SCORE OVER LENGTH SEARCHES

Attached is a score over length search. This search was developed to overcome limitations in most standard search systems which favor large sequences with high scoring, but lesser overall identity over smaller sequences with higher overall identity. This search is especially useful for relatively small nucleic acid or polypeptide target sequences (antisense, fragments, probes, primers, RNAi, epitopes, haptens, etc.) claimed functionally via a form of hybridization and/or identity language and having defined upper and lower polynucleotide and or polypeptide length limits.

The score over length search is performed by first running the query sequence using examiner-specified identity and polynucleotide or protein length limit parameters, and saving 65,000 hits and 0 alignments from each desired database. The resulting output is reformatted using a Microsoft Word macro and is imported into Excel. The summary table data are then sorted by the ratio of score of each hit sequence divided by its length and the accession numbers for all hits below the examiner's desired score over length parameters are deleted. The remaining accession numbers are used to pull the corresponding sequences from the databases into subdatabases enriched for good hits and the query sequence is re-run against these subdatabases to yield the final results.

The score over length cutoff for this search is 70%, length 100 - 7582 nt

Examiner Please Note: This cover sheet should be included when submitting results to be scanned.

No suitable hits were found in Published Applications NA_New
No results from this file included here.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 4, 2006, 17:16:20 ; Search time 2.2 Seconds
(without alignments)
7.487 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 2447
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: UNITARY_AA
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

uni Prot

Searched: 20 seqs, 7241 residues

Total number of hits satisfying chosen parameters: 40

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 20 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=new.seq -DB=unidb -SUFFIX=pto
-OUT=uni.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=unitary_AA -TRANS=human40.cdi -LIST=20 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : unidb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 517 | 21.1 | 538 | 1 | ENW1_HUMAN | HERV-W_7q21.2 prov |
| 2 | 506 | 20.7 | 538 | 1 | ENW1_PANTR | ERV-W1 provirus an |
| 3 | 502 | 20.5 | 538 | 1 | ENW1_GORGO | ERV-W1 provirus an |
| 4 | 502 | 20.5 | 538 | 1 | Q2XSV0_9PRIM | Syncytin 1. Goril |
| 5 | 495 | 20.2 | 538 | 1 | ENW1_PONPY | ERV-W1 provirus an |
| 6 | 495 | 20.2 | 538 | 1 | Q2XSU9_PONPY | Syncytin 1. Pongo |
| 7 | 492 | 20.1 | 538 | 1 | Q2XSU8_HYLSY | Syncytin 1. Hylob |
| 8 | 489 | 20.0 | 538 | 1 | ENW1_HYLPI | ERV-W1 provirus an |
| 9 | 441 | 18.0 | 542 | 1 | Q991W9_9RETR | Recombinant envelo |
| 10 | 291 | 11.9 | 319 | 1 | Q96L59_HUMAN | Envelope protein ( |
| 11 | 290 | 11.9 | 319 | 1 | Q96L60_HUMAN | Envelope protein ( |
| 12 | 290 | 11.9 | 319 | 1 | Q96L61_HUMAN | Envelope protein ( |
| 13 | 177 | 7.2 | 180 | 1 | Q96TB5_HUMAN | Envelope protein ( |
| 14 | 162 | 6.6 | 180 | 1 | Q96TB1_HUMAN | Envelope protein ( |
| 15 | 162 | 6.6 | 180 | 1 | Q96TB4_HUMAN | Envelope protein ( |
| 16 | 157 | 6.4 | 180 | 1 | Q96TB6_HUMAN | Envelope protein ( |
| 17 | 156 | 6.4 | 180 | 1 | Q96TB2_HUMAN | Envelope protein ( |
| 18 | 151 | 6.2 | 180 | 1 | Q96TB7_HUMAN | Envelope protein ( |
| 19 | 150.5 | 6.2 | 179 | 1 | Q96TB3_HUMAN | Envelope protein ( |
| 20 | 147.5 | 6.0 | 179 | 1 | Q96TB8_HUMAN | Envelope protein ( |

ALIGNMENTS

RESULT 1
ENW1_HUMAN
ID ENW1_HUMAN STANDARD; PRT; 538 AA.
AC Q9UQF0; O95244; O95245; Q8NHY7; Q9NRZ2; Q9NZG3;
DT 24-MAY-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-2000, sequence version 1.
DT 07-MAR-2006, entry version 30.
DE HERV-W_7q21.2 provirus ancestral Env polyprotein precursor (Envelope
DE polyprotein) (HERV-7q Envelope protein) (HERV-W envelope protein)
DE (Syncytin) (Syncytin-1) (Enverin) (Env-W) [Contains: Surface protein
DE (SU); Transmembrane protein (TM)].
GN Name=ERVWE1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND VARIANT ASN-307.
RC TISSUE=Placenta;
RX MEDLINE=21985840; PubMed=11990458;
RA Alliel P.M., Perin J.-P., Goudou D., Bitoun M., Robert B., Rieger F.;
RT "The HERV-W/7q family in the human genome. Potential for protein
RT expression and gene regulation.";
RL Cell. Mol. Biol. (Noisy-le-Grand) 48:213-217(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=99099005; PubMed=9882319;
RA Blond J.-L., Beseme F., Duret L., Bouton O., Bedin F., Perron H.,
RA Mandrand B., Mallet F.;
RT "Molecular characterization and placental expression of HERV-W, a new
RT human endogenous retrovirus family.";
RL J. Virol. 73:1175-1185(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, FUNCTION, AND VARIANT
RP ASN-307.
RC TISSUE=Testis;
RX MEDLINE=20155476; PubMed=10693809; DOI=10.1038/35001608;
RA Sha M., Lee X., Li X.-P., Veldman G.M., Finnerty H., Racie L.,
RA LaVallie E., Tang X.-Y., Edouard P., Howes S., Keith J.C. Jr.,
RA McCoy J.M.;
RT "Syncytin is captive retroviral envelope protein involved in human
RT placental morphogenesis.";
RL Nature 403:785-789(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20284713; PubMed=10826480; DOI=10.1089/088922200308738;
RA Voisset C., Bouton O., Bedin F., Duret L., Mandrand B., Mallet F.,
RA Paranhos-Baccala G.;
RT "Chromosomal distribution and coding capacity of the human endogenous
RT retrovirus HERV-W family.";
RL AIDS Res. Hum. Retroviruses 16:731-740(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,

RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA Waterston R.H., Wilson R.K.;
RT "The DNA sequence of human chromosome 7.";
RL Nature 424:157-164(2003).
RN [6]
RP FUNCTION, AND TISSUE SPECIFICITY.
RX PubMed=10708449; DOI=10.1128/JVI.74.7.3321-3329.2000;
RA Blond J.-L., Lavillette D., Cheynet V., Bouton O., Oriol G.,
RA Chapel-Fernandes S., Mandrand B., Mallet F., Cosset F.-L.;
RT "An envelope glycoprotein of the human endogenous retrovirus HERV-W is
RT expressed in the human placenta and fuses cells expressing the type D
RT mammalian retrovirus receptor.";
RL J. Virol. 74:3321-3329(2000).
RN [7]
RP FUNCTION.
RX PubMed=11238877; DOI=10.1128/JVI.75.7.3488-3489.2001;
RA An D.S., Xie Y.-M., Chen I.S.Y.;
RT "Envelope gene of the human endogenous retrovirus HERV-W encodes a
RT functional retrovirus envelope.";
RL J. Virol. 75:3488-3489(2001).
RN [8]
RP FUNCTION.
RX PubMed=11531410; DOI=10.1006/viro.2001.1045;
RA Perron H., Jouvin-Marche E., Michel M., Ounanian-Paraz A., Camelo S.,
RA Dumon A., Jolivet-Reynaud C., Marcel F., Souillet Y., Borel E.,
RA Gebuhrer L., Santoro L., Marcel S., Seigneurin J.M., Marche P.N.,
RA Lafon M.;
RT "Multiple sclerosis retrovirus particles and recombinant envelope
RT trigger an abnormal immune response in vitro, by inducing polyclonal
RT Vbeta16 T-lymphocyte activation.";
RL Virology 287:321-332(2001).
RN [9]
RP FUNCTION.
RX PubMed=12050356; DOI=10.1128/JVI.76.13.6442-6452.2002;
RA Lavillette D., Marin M., Ruggieri A., Mallet F., Cosset F.-L.,
RA Kabat D.;
RT "The envelope glycoprotein of human endogenous retrovirus type W uses
RT a divergent family of amino acid transporters/cell surface
RT receptors.";
RL J. Virol. 76:6442-6452(2002).
RN [10]
RP FUNCTION.
RX MEDLINE=22946171; PubMed=14557543; DOI=10.1073/pnas.2132646100;
RA Blaise S., de Parseval N., Benit L., Heidmann T.;
RT "Genomewide screening for fusogenic human endogenous retrovirus
RT envelopes identifies syncytin 2, a gene conserved on primate
RT evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:13013-13018(2003).
RN [11]
RP FUNCTION.
RX PubMed=12664292; DOI=10.1007/s00705-002-0960-x;
RA Ponferrada V.G., Mauck B.S., Wooley D.P.;
RT "The envelope glycoprotein of human endogenous retrovirus HERV-W
RT induces cellular resistance to spleen necrosis virus.";
RL Arch. Virol. 148:659-675(2003).
RN [12]
RP TISSUE SPECIFICITY.
RX MEDLINE=22850809; PubMed=12970426;
RX DOI=10.1128/JVI.77.19.10414-10422.2003;
RA de Parseval N., Lazar V., Casella J.-F., Benit L., Heidmann T.;
RT "Survey of human genes of retroviral origin: identification and
RT transcriptome of the genes with coding capacity for complete envelope
RT proteins.";
RL J. Virol. 77:10414-10422(2003).
RN [13]
RP DEVELOPMENTAL STAGE.
RX PubMed=12620933; DOI=10.1095/biolreprod.102.013078;
RA Smallwood A., Papageorghiou A., Nicolaides K., Alley M.K.R., Jim A.,
RA Nargund G., Ojha K., Campbell S., Banerjee S.;
RT "Temporal regulation of the expression of syncytin (HERV-W),
RT maternally imprinted PEG10, and SGCE in human placenta.";
RL Biol. Reprod. 69:286-293(2003).
RN [14]
RP INVOLVEMENT IN MULTIPLE SCLEROSIS, AND TISSUE SPECIFICITY.
RX PubMed=15452578; DOI=10.1038/nn1319;
RA Antony J.M., Van Marle G., Opii W., Butterfield D.A., Mallet F.,
RA Yong V.W., Wallace J.L., Deacon R.M., Warren K., Power C.;
RT "Human endogenous retrovirus glycoprotein-mediated induction of redox
RT reactants causes oligodendrocyte death and demyelination.";
RL Nat. Neurosci. 7:1088-1095(2004).
RN [15]
RP IDENTIFICATION.
RX MEDLINE=99052087; PubMed=9835022; DOI=10.1016/S0764-4469(99)80026-2;
RA Alliel P.M., Perin J.P., Pierig R., Nussbaum J.L., Menard A.,
RA Rieger F.;
RT "Endogenous retroviruses and multiple sclerosis. II. HERV-7q.";
RL C. R. Acad. Sci. III, Sci. Vie 321:857-863(1998).
RN [16]
RP VARIANTS ALA-129; GLN-138; ASN-307 AND PHE-477.
RX PubMed=14757826; DOI=10.1073/pnas.0305763101;
RA Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B.,
RA Lucotte G., Duret L., Mandrand B.;
RT "The endogenous retroviral locus ERVWE1 is a bona fide gene involved
RT in hominoid placental physiology.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:1731-1736(2004).
CC -!- FUNCTION: Retroviral envelope proteins mediate receptor
CC recognition and membrane fusion during early infection. Endogenous
CC envelope proteins may have kept, lost or modified their original
CC function during evolution. This endogenous envelope protein has
CC retained its original fusogenic properties. It can make
CC pseudotypes with HIV-1 virions and confer infectivity. It
CC participates to trophoblast fusion during placenta morphogenesis.
CC In vitro it fuses cells expressing the type D mammalian retrovirus
CC receptors SLC1A4 and SLC1A5. It can induce cellular resistance to
CC spleen necrosis virus in vitro.
CC -!- FUNCTION: SU mediates receptor recognition (By similarity).
CC -!- FUNCTION: TM anchors the envelope heterodimer to the viral
CC membrane through one transmembrane domain. The other transmembrane
CC domain, called fusion peptide, mediates fusion of the viral
CC membrane with the target cell membrane (By similarity).
CC -!- SUBUNIT: The surface (SU) and transmembrane (TM) proteins form a
CC heterodimer. SU and TM are attached by noncovalent interactions or
CC by a labile interchain disulfide bond (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed at higher level in placenta.
CC Expressed at lower level in adrenal tissue, bone marrow, breast,
CC colon, kidney, ovary, prostate, skin, spleen, testis, thymus,
CC thyroid and trachea. Low expression is observed in the brain. Both
CC mRNA and protein levels are significantly increased in the brain
CC of individuals with multiple sclerosis, particularly in astrocytes
CC and microglia.
CC -!- DEVELOPMENTAL STAGE: In placenta, detected at higher level during
CC early pregnancy and at lower level during late pregnancy.
CC -!- DOMAIN: The CKS-17 immunosuppressive domain is present in many
CC retroviral envelope proteins. As a synthetic peptide, it inhibits
CC immune function in vitro and in vivo (By similarity).
CC -!- PTM: Specific enzymatic cleavages in vivo yield the mature SU and
CC TM proteins (By similarity).

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.173 | Length: | 538 |
| Score: | 517.00 | Matches: | 517 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x ENW1_HUMAN (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
                 |||   ||||||||||||||||||||      |||      |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20
```

```
Qy      5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db        21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy      5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db        41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db       101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db       121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db       141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db       161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db       181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db       201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db       221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db       241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db       261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db       281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db       301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db       321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db       341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db       361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db       381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db       401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db       421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db       441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db       481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy      7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db       501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy      7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db       521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

```
RESULT 2
ENW1_PANTR
ID   ENW1_PANTR     STANDARD;      PRT;   538 AA.
AC   P61563;
DT   24-MAY-2004, integrated into UniProtKB/Swiss-Prot.
DT   24-MAY-2004, sequence version 1.
DT   07-FEB-2006, entry version 16.
DE   ERV-W1 provirus ancestral Env polyprotein precursor (Envelope
DE   polyprotein) (ERVWE1 envelope protein) (Syncytin) [Contains: Surface
DE   protein (SU); Transmembrane protein (TM)].
GN   Name=ERVWE1;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   PubMed=14757826; DOI=10.1073/pnas.0305763101;
RA   Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B.,
RA   Lucotte G., Duret L., Mandrand B.;
RT   "The endogenous retroviral locus ERVWE1 is a bona fide gene involved
RT   in hominoid placental physiology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:1731-1736(2004).
CC   -!- FUNCTION: Retroviral envelope proteins mediate receptor
CC       recognition and membrane fusion during early infection. Endogenous
CC       envelope proteins may have kept, lost or modified their original
CC       function during evolution. This endogenous envelope protein has
CC       retained its original fusogenic properties. In vitro it fuses
CC       cells expressing the type D mammalian retrovirus receptors SLC1A4
CC       and SLC1A5.
CC   -!- FUNCTION: SU mediates receptor recognition (By similarity).
CC   -!- FUNCTION: TM anchors the envelope heterodimer to the viral
CC       membrane through one transmembrane domain. The other transmembrane
CC       domain, called fusion peptide, mediates fusion of the viral
CC       membrane with the target cell membrane (By similarity).
CC   -!- SUBUNIT: The surface (SU) and transmembrane (TM) proteins form a
CC       heterodimer. SU and TM are attached by noncovalent interactions or
CC       by a labile interchain disulfide bond (By similarity).
CC   -!- DOMAIN: The CKS-17 immunosuppressive domain is present in many
CC       retroviral envelope proteins. As a synthetic peptide, it inhibits
```

```
CC       immune function in vitro and in vivo (By similarity).
CC   -!- PTM: Specific enzymatic cleavages in vivo yield the mature SU and
CC       TM proteins (By similarity).
CC   -!- PTM: The CXXC motif is highly conserved across a broad range of
CC       retroviral envelope proteins. It is thought to participate to the
CC       formation of a labile disulfide bond possibly with the CX6CC motif
CC       present in the transmembrane protein (By similarity).
CC   -!- MISCELLANEOUS: Ortholog of the human HERV-W_7q21.2 envelope
CC       protein.
CC   -!- SIMILARITY: Belongs to the gamma type-C retroviral envelope
CC       protein family. ERV class-I W subfamily.
CC   -!- CAUTION: CKS-17 sequence does not match the minimal active
CC       concensus.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY101586; AAM68165.1; -; Genomic_DNA.
DR   EMBL; AY101587; AAM68166.1; -; Genomic_DNA.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Cleavage on pair of basic residues; Envelope protein; ERV;
KW   Glycoprotein; Membrane; Polyprotein; Signal; Transmembrane;
KW   Transposable element.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    538       ERV-W1 provirus ancestral Env
FT                                 polyprotein.
FT                                 /FTId=PRO_0000008491.
FT   CHAIN        21    317       Surface protein (By similarity).
FT                                 /FTId=PRO_0000008492.
FT   CHAIN       318    538       Transmembrane protein (By similarity).
FT                                 /FTId=PRO_0000008493.
FT   TRANSMEM    320    340       Fusion peptide (Potential).
FT   TRANSMEM    444    464       Potential.
FT   MOTIF       186    189       CXXC (By similarity).
FT   MOTIF       380    396       CKS-17 (By similarity).
FT   MOTIF       397    405       CX6CC (By similarity).
FT   SITE        317    318       Cleavage (By similarity).
FT   CARBOHYD    169    169       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    208    208       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    214    214       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    234    234       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    242    242       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    281    281       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    409    409       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   538 AA;  59981 MW;  A5DB5C51EF33AB0C CRC64;

Alignment Scores:
Pred. No.:              0.195          Length:         538
Score:                  506.00         Matches:        506
Percent Similarity:     94.1%          Conservative:   0
Best Local Similarity:  94.1%          Mismatches:     32
Query Match:            20.7%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x ENW1_PANTR (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysIleIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerAlaSerArgProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440
```

```
Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              |||||||||      ||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleLeuThrAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 3
ENW1_GORGO
ID   ENW1_GORGO     STANDARD;      PRT;   538 AA.
AC   P61561;
DT   24-MAY-2004, integrated into UniProtKB/Swiss-Prot.
DT   24-MAY-2004, sequence version 1.
DT   07-MAR-2006, entry version 17.
DE   ERV-W1 provirus ancestral Env polyprotein precursor (Envelope
DE   polyprotein) (ERVWE1 envelope protein) (Syncytin) (Syncytin-1)
DE   [Contains: Surface protein (SU); Transmembrane protein (TM)].
GN   Name=ERVWE1;
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   PubMed=14757826; DOI=10.1073/pnas.0305763101;
RA   Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B.,
RA   Lucotte G., Duret L., Mandrand B.;
RT   "The endogenous retroviral locus ERVWE1 is a bona fide gene involved
RT   in hominoid placental physiology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:1731-1736(2004).
CC   -!- FUNCTION: Retroviral envelope proteins mediate receptor
CC       recognition and membrane fusion during early infection. Endogenous
CC       envelope proteins may have kept, lost or modified their original
CC       function during evolution. This endogenous envelope protein has
CC       retained its original fusogenic properties. In vitro it fuses
CC       cells expressing the type D mammalian retrovirus receptors SLC1A4
CC       and SLC1A5.
CC   -!- FUNCTION: SU mediates receptor recognition (By similarity).
CC   -!- FUNCTION: TM anchors the envelope heterodimer to the viral
CC       membrane through one transmembrane domain. The other transmembrane
CC       domain, called fusion peptide, mediates fusion of the viral
CC       membrane with the target cell membrane (By similarity).
CC   -!- SUBUNIT: The surface (SU) and transmembrane (TM) proteins form a
CC       heterodimer. SU and TM are attached by noncovalent interactions or
CC       by a labile interchain disulfide bond (By similarity).
CC   -!- DOMAIN: The CKS-17 immunosuppressive domain is present in many
CC       retroviral envelope proteins. As a synthetic peptide, it inhibits
CC       immune function in vitro and in vivo (By similarity).
CC   -!- PTM: Specific enzymatic cleavages in vivo yield the mature SU and
CC       TM proteins (By similarity).
CC   -!- PTM: The CXXC motif is highly conserved across a broad range of
CC       retroviral envelope proteins. It is thought to participate to the
CC       formation of a labile disulfide bond possibly with the CX6CC motif
CC       present in the transmembrane protein (By similarity).
CC   -!- MISCELLANEOUS: Ortholog of the human HERV-W_7q21.2 envelope
CC       protein.
CC   -!- SIMILARITY: Belongs to the gamma type-C retroviral envelope
CC       protein family. ERV class-I W subfamily.
CC   -!- CAUTION: CKS-17 sequence does not match the minimal active
CC       consensus.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY101588; AAM68167.1; -; Genomic_DNA.
DR   EMBL; AY101589; AAM68168.1; -; Genomic_DNA.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Cleavage on pair of basic residues; Envelope protein; ERV;
KW   Glycoprotein; Membrane; Polyprotein; Signal; Transmembrane;
KW   Transposable element.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    538       ERV-W1 provirus ancestral Env
FT                                 polyprotein.
FT                                 /FTId=PRO_0000008482.
FT   CHAIN        21    317       Surface protein (By similarity).
FT                                 /FTId=PRO_0000008483.
FT   CHAIN       318    538       Transmembrane protein (By similarity).
FT                                 /FTId=PRO_0000008484.
FT   TRANSMEM    320    340       Fusion peptide (Potential).
FT   TRANSMEM    444    464       Potential.
FT   MOTIF       186    189       CXXC (By similarity).
FT   MOTIF       380    396       CKS-17 (By similarity).
FT   MOTIF       397    405       CX6CC (By similarity).
FT   SITE        317    318       Cleavage (By similarity).
FT   CARBOHYD    169    169       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    208    208       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    214    214       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    234    234       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    281    281       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    409    409       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   538 AA;  59939 MW;  95D4393FEDE196F3 CRC64;

Alignment Scores:
Pred. No.:                0.204          Length:          538
Score:                    502.00         Matches:         502
Percent Similarity:       93.3%          Conservative:    0
Best Local Similarity:    93.3%          Mismatches:      36
Query Match:              20.5%          Indels:          0
DB:                       1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x ENW1_GORGO (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||||||||   |||       |||       |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIleLeuLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||   ||||||||||||||||||||||||   ||||||      ||||||||||||
Db         41 MetArgArgProGlyAsnIleAspAlaProSerHisArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140
```

```
Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAspPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGluTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
```

```
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysCysThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerThrAlaGlnProLeuArgProAsnSerAlaGlySerSer 538

RESULT 4
Q2XSV0_9PRIM
ID   Q2XSV0_9PRIM   PRELIMINARY;   PRT;   538 AA.
AC   Q2XSV0;
DT   20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   20-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 3.
DE   Syncytin 1.
OS   Gorilla gorilla (gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Gorilla.
OX   NCBI_TaxID=9593;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Caceres M., Thomas J.W.;
RT   "The gene of retroviral origin syncytin 1 is specific to hominoids and
RT   was inactivated in Old World monkeys.";
RL   Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; DQ256474; ABB73024.1; -; Genomic_DNA.
SQ   SEQUENCE   538 AA;  59913 MW;  B31A55D9238D7497 CRC64;

Alignment Scores:
Pred. No.:              0.204          Length:         538
Score:                  502.00         Matches:        502
Percent Similarity:     93.3%          Conservative:   0
Best Local Similarity:  93.3%          Mismatches:     36
Query Match:            20.5%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q2XSV0_9PRIM (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIleLeuLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetArgArgProGlyAsnIleAspAlaProSer***ArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
```

```
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||   |||   ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAspPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGluTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              ||||||||||||||   ||||||||||||||||||||||||   |||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysCysThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerThrAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 5
ENW1_PONPY
ID   ENW1_PONPY     STANDARD;      PRT;   538 AA.
AC   P61564;
DT   24-MAY-2004, integrated into UniProtKB/Swiss-Prot.
DT   24-MAY-2004, sequence version 1.
DT   07-FEB-2006, entry version 16.
DE   ERV-W1 provirus ancestral Env polyprotein precursor (Envelope
DE   polyprotein) (ERVWE1 envelope protein) (Syncytin) [Contains: Surface
DE   protein (SU); Transmembrane protein (TM)].
GN   Name=ERVWE1;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   PubMed=14757826; DOI=10.1073/pnas.0305763101;
RA   Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B.,
RA   Lucotte G., Duret L., Mandrand B.;
RT   "The endogenous retroviral locus ERVWE1 is a bona fide gene involved
RT   in hominoid placental physiology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:1731-1736(2004).
CC   -!- FUNCTION: Retroviral envelope proteins mediate receptor
CC       recognition and membrane fusion during early infection. Endogenous
CC       envelope proteins may have kept, lost or modified their original
CC       function during evolution. This endogenous envelope protein has
CC       retained its original fusogenic properties. In vitro it fuses
CC       cells expressing the type D mammalian retrovirus receptors SLC1A4
CC       and SLC1A5.
CC   -!- FUNCTION: SU mediates receptor recognition (By similarity).
CC   -!- FUNCTION: TM anchors the envelope heterodimer to the viral
CC       membrane through one transmembrane domain. The other transmembrane
CC       domain, called fusion peptide, mediates fusion of the viral
CC       membrane with the target cell membrane (By similarity).
CC   -!- SUBUNIT: The surface (SU) and transmembrane (TM) proteins form a
CC       heterodimer. SU and TM are attached by noncovalent interactions or
CC       by a labile interchain disulfide bond (By similarity).
CC   -!- DOMAIN: The CKS-17 immunosuppressive domain is present in many
CC       retroviral envelope proteins. As a synthetic peptide, it inhibits
CC       immune function in vitro and in vivo (By similarity).
CC   -!- PTM: Specific enzymatic cleavages in vivo yield the mature SU and
CC       TM proteins (By similarity).
CC   -!- PTM: The CXXC motif is highly conserved across a broad range of
CC       retroviral envelope proteins. It is thought to participate to the
CC       formation of a labile disulfide bond possibly with the CX6CC motif
CC       present in the transmembrane protein (By similarity).
CC   -!- MISCELLANEOUS: Ortholog of the human HERV-W_7q21.2 envelope
CC       protein.
CC   -!- SIMILARITY: Belongs to the gamma type-C retroviral envelope
CC       protein family. ERV class-I W subfamily.
CC   -!- CAUTION: CKS-17 sequence does not match the minimal active
CC       concensus.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY101590; AAM68169.1; -; Genomic_DNA.
DR   EMBL; AY101591; AAM68170.1; -; Genomic_DNA.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Cleavage on pair of basic residues; Envelope protein; ERV;
KW   Glycoprotein; Membrane; Polyprotein; Signal; Transmembrane;
KW   Transposable element.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    538       ERV-W1 provirus ancestral Env
```

```
FT                                   polyprotein.
FT                                   /FTId=PRO_0000008494.
FT   CHAIN        21    317          Surface protein (By similarity).
FT                                   /FTId=PRO_0000008495.
FT   CHAIN       318    538          Transmembrane protein (By similarity).
FT                                   /FTId=PRO_0000008496.
FT   TRANSMEM    320    340          Fusion peptide (Potential).
FT   TRANSMEM    444    464          Potential.
FT   MOTIF       186    189          CXXC (By similarity).
FT   MOTIF       380    396          CKS-17 (By similarity).
FT   MOTIF       397    405          CX6CC (By similarity).
FT   SITE        317    318          Cleavage (By similarity).
FT   CARBOHYD    169    169          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    208    208          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    214    214          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    234    234          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    242    242          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    245    245          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    281    281          N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    409    409          N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   538 AA;  59909 MW;  AA17EE08E6D62A99 CRC64;

Alignment Scores:
Pred. No.:                0.221          Length:        538
Score:                    495.00         Matches:       495
Percent Similarity:       92.0%          Conservative:  0
Best Local Similarity:    92.0%          Mismatches:    43
Query Match:              20.2%          Indels:        0
DB:                       1              Gaps:          0

US-10-717-580-11-COPY (1-7582) x ENW1_PONPY (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||  ||||||||||||||||||||      |||     |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||      ||||||||||||||||||||||||||||||||      ||||||||||||
Db         41 MetHisHisProGlyAsnIleAspAlaProSerTyrArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisGlyAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||   ||||||||||||||||||||||||||||||||||||||   ||||||
Db        121 GlnAspGluAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlnValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||   ||||||||||   |||   ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuAlaPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrAsnThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ValValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValMetAlaAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerIleAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||||   ||||||||||||||   ||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnIleGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleLeuLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrHisArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              |||||||||   |||   ||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspTrpProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerThrAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 6
Q2XSU9_PONPY
ID   Q2XSU9_PONPY   PRELIMINARY;   PRT;   538 AA.
AC   Q2XSU9;
DT   20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   20-DEC-2005, sequence version 1.
```

```
DT   07-FEB-2006, entry version 3.
DE   Syncytin 1.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Caceres M., Thomas J.W.;
RT   "The gene of retroviral origin syncytin 1 is specific to hominoids and
RT   was inactivated in Old World monkeys.";
RL   Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; DQ256475; ABB73025.1; -; Genomic_DNA.
SQ   SEQUENCE   538 AA;  59905 MW;  EE11C048E13956EA CRC64;

Alignment Scores:
Pred. No.:              0.221          Length:          538
Score:                  495.00         Matches:         495
Percent Similarity:     92.0%          Conservative:    0
Best Local Similarity:  92.0%          Mismatches:      43
Query Match:            20.2%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x Q2XSU9_PONPY (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPhe***ValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheValTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetHisHisProGlyAsnIleAspAlaProSerTyrArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisGlyAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGluAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlnValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuAlaPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrAsnThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 ValValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValMetAlaAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerIleAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnIleGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 PheSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrHisArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspTrpProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerThrAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 7
Q2XSU8_HYLSY
ID   Q2XSU8_HYLSY   PRELIMINARY;   PRT;   538 AA.
AC   Q2XSU8;
DT   20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   20-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 3.
DE   Syncytin 1.
```

```
OS   Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Hylobatidae; Symphalangus.
OX   NCBI_TaxID=9590;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Caceres M., Thomas J.W.;
RT   "The gene of retroviral origin syncytin 1 is specific to hominoids and
RT   was inactivated in Old World monkeys.";
RL   Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; DQ256476; ABB73026.1; -; Genomic_DNA.
SQ   SEQUENCE   538 AA;  60020 MW;  1A56BC75B8C907E7 CRC64;

Alignment Scores:
Pred. No.:              0.228          Length:         538
Score:                  492.00         Matches:        492
Percent Similarity:     91.4%          Conservative:   0
Best Local Similarity:  91.4%          Mismatches:     46
Query Match:            20.1%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q2XSU8_HYLSY (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnHisProGlyAsnIleAspAlaProSerTyrArgSerPheSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrGlnTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAsnGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlnValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuAspPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IlePheCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValValSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValMetGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerArgGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaIleLysLeuGlnMetGluProLysMetGluSerLysThrLysThrTyrHisArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 SerLeuAspTrpProValSerProArgSerAspValAsnAspIleLysGlyIleProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerThrAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 8
ENW1_HYLPI
ID   ENW1_HYLPI     STANDARD;      PRT;   538 AA.
AC   P61562;
DT   24-MAY-2004, integrated into UniProtKB/Swiss-Prot.
DT   24-MAY-2004, sequence version 1.
DT   07-FEB-2006, entry version 16.
DE   ERV-W1 provirus ancestral Env polyprotein precursor (Envelope
DE   polyprotein) (ERVWE1 envelope protein) (Syncytin) [Contains: Surface
DE   protein (SU); Transmembrane protein (TM)].
```

```
GN   Name=ERVWE1;
OS   Hylobates pileatus (Pileated gibbon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Hylobatidae; Hylobates.
OX   NCBI_TaxID=9589;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   PubMed=14757826; DOI=10.1073/pnas.0305763101;
RA   Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B.,
RA   Lucotte G., Duret L., Mandrand B.;
RT   "The endogenous retroviral locus ERVWE1 is a bona fide gene involved
RT   in hominoid placental physiology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:1731-1736(2004).
CC   -!- FUNCTION: Retroviral envelope proteins mediate receptor
CC       recognition and membrane fusion during early infection. Endogenous
CC       envelope proteins may have kept, lost or modified their original
CC       function during evolution. This endogenous envelope protein has
CC       retained its original fusogenic properties. In vitro it fuses
CC       cells expressing the type D mammalian retrovirus receptors SLC1A4
CC       and SLC1A5.
CC   -!- FUNCTION: SU mediates receptor recognition (By similarity).
CC   -!- FUNCTION: TM anchors the envelope heterodimer to the viral
CC       membrane through one transmembrane domain. The other transmembrane
CC       domain, called fusion peptide, mediates fusion of the viral
CC       membrane with the target cell membrane (By similarity).
CC   -!- SUBUNIT: The surface (SU) and transmembrane (TM) proteins form a
CC       heterodimer. SU and TM are attached by noncovalent interactions or
CC       by a labile interchain disulfide bond (By similarity).
CC   -!- DOMAIN: The CKS-17 immunosuppressive domain is present in many
CC       retroviral envelope proteins. As a synthetic peptide, it inhibits
CC       immune function in vitro and in vivo (By similarity).
CC   -!- PTM: Specific enzymatic cleavages in vivo yield the mature SU and
CC       TM proteins (By similarity).
CC   -!- PTM: The CXXC motif is highly conserved across a broad range of
CC       retroviral envelope proteins. It is thought to participate to the
CC       formation of a labile disulfide bond possibly with the CX6CC motif
CC       present in the transmembrane protein (By similarity).
CC   -!- MISCELLANEOUS: Ortholog of the human HERV-W_7q21.2 envelope
CC       protein.
CC   -!- SIMILARITY: Belongs to the gamma type-C retroviral envelope
CC       protein family. ERV class-I W subfamily.
CC   -!- CAUTION: CKS-17 sequence does not match the minimal active
CC       concensus.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY101592; AAM68171.1; -; Genomic_DNA.
DR   EMBL; AY101593; AAM68172.1; -; Genomic_DNA.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Cleavage on pair of basic residues; Envelope protein; ERV;
KW   Glycoprotein; Membrane; Polyprotein; Signal; Transmembrane;
KW   Transposable element.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    538       ERV-W1 provirus ancestral Env
FT                                 polyprotein.
FT                                 /FTId=PRO_0000008488.
FT   CHAIN        21    317       Surface protein (By similarity).
FT                                 /FTId=PRO_0000008489.
FT   CHAIN       318    538       Transmembrane protein (By similarity).
FT                                 /FTId=PRO_0000008490.
FT   TRANSMEM    320    340       Fusion peptide (Potential).
FT   TRANSMEM    444    464       Potential.
FT   MOTIF       186    189       CXXC (By similarity).
FT   MOTIF       380    396       CKS-17 (By similarity).
FT   MOTIF       397    405       CX6CC (By similarity).
FT   SITE        317    318       Cleavage (By similarity).
FT   CARBOHYD    169    169       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    208    208       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    214    214       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    234    234       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    242    242       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    281    281       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    409    409       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   538 AA;  60100 MW;  E833A5AF7AF66BD7 CRC64;

Alignment Scores:
Pred. No.:              0.236          Length:         538
Score:                  489.00         Matches:        489
Percent Similarity:     90.9%          Conservative:   0
Best Local Similarity:  90.9%          Mismatches:     49
Query Match:            20.0%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x ENW1_HYLPI (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
                ||| |||||||||||||||||||||      |||      |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
                |||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrGlnArgProGlyAsnIleAspAlaProLeuTyrArgSerPheSerLysGlySerPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
                |||||||||||||||||  |||||||||  ||||||  ||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrTyrMetProArgThrCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
                |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrGlnTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
                |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlnValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
                |||||||||||||||||||  ||||||||||||  |||  ||||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuAspPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
                |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
                ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
                |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IlePheCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
```

```
                ||||||
```

```
Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||      ||||||  ||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||      |||      ||||||  ||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||  ||||||||||||||||  ||||||||||||||||||  |||
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                ||||||||||||||||||  ||||||||  ||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||  ||||||||||||||||||||  ||||||||||||||||||  |||
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||  ||||||||||||  ||||||||||||  ||||||  |||||||||
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||  ||||||||||||  ||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpThrLeuProPheLeuGlyProLeuAlaAlaIleIlePhe 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||      ||||||  ||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnPheLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAA------------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              ||||||||||||            ||||||||||||||  |||||||||  ||||||
Db        481 GluAlaValLysLeuGlnIleValLeuGlnMetGluProGlnMetGlnSerMetThrLys 500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              |||||||||  ||||||||||||||||      ||||||||||||||||||
Db        501 IleTyrArgGlyProLeuAspArgProAlaArgLeuCysSerAspValAsnAspIleGlu 520

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
                ||||||||||||||||||||  ||||||||||||||||      ||||||  |||
Db        521 ValThrProProGluGluIleSerThrAlaGlnProLeuLeuHisSerAsnSerValGly 540

Qy       7189 AGCAGT 7194
              ||||||
Db        541 SerSer 542

RESULT 10
Q96L59_HUMAN
ID   Q96L59_HUMAN   PRELIMINARY;   PRT;   319 AA.
AC   Q96L59;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 9.
DE   Envelope protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jern P., Karlsson D., Blomberg J.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY050301; AAL11495.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000197604; Homo sapiens.
DR   GO; GO:0019031; C:viral envelope; IEA.
KW   Envelope protein.
FT   NON_TER     319     319
SQ   SEQUENCE   319 AA;  35784 MW;  7D288CCCD1B68999 CRC64;

Alignment Scores:
Pred. No.:              3.49          Length:          319
Score:                  291.00        Matches:         291
Percent Similarity:     91.2%         Conservative:    0
Best Local Similarity:  91.2%         Mismatches:      28
Query Match:            11.9%         Indels:          0
DB:                     1             Gaps:            0

US-10-717-580-11-COPY (1-7582) x Q96L59_HUMAN (1-319)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||  ||||||||||||||||||||      |||      |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrAlaLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||  |||  ||||||||||||  |||  ||||||  |||
Db        181 GlnAsnProThrAsnSerTrpMetCysLeuProLeuHisPheArgProTyrIleSerThr 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleIleThrThrSerValLeuVal 220
```

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuAlaSerAsnValGluIleIleHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnSerPheSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 319

RESULT 11
Q96L60_HUMAN
ID   Q96L60_HUMAN   PRELIMINARY;   PRT;   319 AA.
AC   Q96L60;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 9.
DE   Envelope protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jern P., Karlsson D., Blomberg J.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY050300; AAL11494.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000197604; Homo sapiens.
DR   LinkHub; Q96L60; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
KW   Envelope protein.
FT   NON_TER     319    319
SQ   SEQUENCE   319 AA;  35764 MW;  67822669D1B6899D CRC64;

Alignment Scores:
Pred. No.:              3.53          Length:         319
Score:                  290.00        Matches:        290
Percent Similarity:     90.9%         Conservative:   0
Best Local Similarity:  90.9%         Mismatches:     29
Query Match:            11.9%         Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96L60_HUMAN (1-319)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys*** 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrAlaLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnSerTrpMetCysLeuProLeuHisPheArgProTyrIleSerThr 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleIleThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuAlaSerAsnValGluIleIleHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrAsnSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnSerPheSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 319

RESULT 12
Q96L61_HUMAN
ID   Q96L61_HUMAN   PRELIMINARY;   PRT;   319 AA.
AC   Q96L61;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 9.
DE   Envelope protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jern P., Karlsson D., Blomberg J.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AY050299; AAL11493.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000197604; Homo sapiens.
```

```
DR   GO; GO:0019031; C:viral envelope; IEA.
KW   Envelope protein.
FT   NON_TER     319     319
SQ   SEQUENCE   319 AA;  35655 MW;  7D3397CCCAAD8999 CRC64;

Alignment Scores:
Pred. No.:              3.53          Length:         319
Score:                  290.00        Matches:        290
Percent Similarity:     90.9%         Conservative:   0
Best Local Similarity:  90.9%         Mismatches:     29
Query Match:            11.9%         Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96L61_HUMAN (1-319)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrAlaLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnSerTrpMetCysLeuProLeuHisPheArgProTyrIleSerThr 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleIleThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuAlaSerAsnValGluIleIleHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrAsnSerGlnCysIleArgGlyValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
```

```
Db        281 AsnSerPheSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 319

RESULT 13
Q96TB5_HUMAN
ID   Q96TB5_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB5;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB051004; BAB47558.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000197604; Homo sapiens.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1       1
FT   NON_TER     180     180
SQ   SEQUENCE   180 AA;  19649 MW;  8D2F33813F16090B CRC64;

Alignment Scores:
Pred. No.:              20            Length:         180
Score:                  177.00        Matches:        177
Percent Similarity:     98.3%         Conservative:   0
Best Local Similarity:  98.3%         Mismatches:     3
Query Match:            7.2%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB5_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         21 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db         41 ThrAsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSer 80
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         81 MetCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerTyrValIleSerLysProArgAsnLysArgValProIleLeuProPheValIleGly 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        161 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeu 180

RESULT 14
Q96TB1_HUMAN
ID   Q96TB1_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB1;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 12.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB051010; BAB47562.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000196702; Homo sapiens.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     180    180
SQ   SEQUENCE   180 AA;  19759 MW;  C4D3C9F72388A0C3 CRC64;

Alignment Scores:
Pred. No.:              23.1          Length:         180
Score:                  162.00        Matches:        162
Percent Similarity:     90.0%         Conservative:   0
Best Local Similarity:  90.0%         Mismatches:     18
Query Match:            6.6%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB1_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         21 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||   ||||||||||||||||||||||||||||||   ||||||||||||||||||
Db         41 ThrSerSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuProSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGluSer 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         81 MetCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                 |||      ||||||  ||||||||||||||||||||||||||||||||||||
Db        101 AsnHisValValProLysProHisAsnLysArgValProIleLeuProPheValIleArg 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        121 AlaGlyValLeuGlyArgLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||   ||||||||||||||   |||   ||||||||||||
Db        141 TyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGlnValThrAspSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        161 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeu 180

RESULT 15
Q96TB4_HUMAN
ID   Q96TB4_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB4;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 12.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB051007; BAB47559.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000198497; Homo sapiens.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     180    180
```

```
SQ   SEQUENCE   180 AA;  20048 MW;  F2B7B2FAF7CE40C3 CRC64;

Alignment Scores:
Pred. No.:              23.1          Length:         180
Score:                  162.00        Matches:        162
Percent Similarity:     90.0%         Conservative:   0
Best Local Similarity:  90.0%         Mismatches:     18
Query Match:            6.6%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB4_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerIleLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                ||||||||||||||||||||||||||||  ||||||||||||||||||      |||
Db         21 LeuGluIleThrHisThrSerAsnLeuThrArgValLysPheSerAsnThrIleAspThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrAsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||  ||||||||||||  |||||||||
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlyPheSerGluSer 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||  ||||||  ||||||||||||||||||||||||
Db         81 MetCysPheLeuSerPheLeuValHisProMetThrIleTyrThrGluGlnAspLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AsnTyrValValProLysProArgAsnLysArgValProIleLeuProPheValIleArg 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||  ||||||  ||||||||||||  |||  ||||||||||||||||||||||||
Db        121 AlaGlyMetLeuGlyArgLeuGlyThrGlyMetGlySerIleThrThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        161 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeu 180

RESULT 16
Q96TB6_HUMAN
ID   Q96TB6_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB6;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB051000; BAB47557.1; -; Genomic_DNA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     180    180
SQ   SEQUENCE   180 AA;  19656 MW;  EF8DFCBD4CAEAF3D CRC64;

Alignment Scores:
Pred. No.:              24.2          Length:         180
Score:                  157.00        Matches:        157
Percent Similarity:     87.2%         Conservative:   0
Best Local Similarity:  87.2%         Mismatches:     23
Query Match:            6.4%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB6_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                ||||||||||||||||||||||||||||||||||||||||||||||||      |||
Db         21 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrAsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||  |||
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlySer 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||  ||||||||||||||  |||||||||  ||||||||||||||||||||||||||
Db         81 MetTrpPheLeuSerPheLeuAlaProProMetIleIleTyrThrGluGlnAspLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              |||||||||  |||  ||||||||||||||||||||  ||||||||||||||||||||||
Db        101 AsnTyrValIleProLysSerHisAsnLysArgValProIleLeuSerPheValIleGly 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||      ||||||  ||||||||||||||  ||||||||||||||||||
Db        121 AlaGlyValLeuAspGlyLeuGlyAlaGlyIleGlyGlyIleLysThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||      |||      |||||||||
Db        141 TyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetAspTrpValThrAsnSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        161 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValProGlnAsnArgArgAlaLeu 180

RESULT 17
Q96TB2_HUMAN
ID   Q96TB2_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB2;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
```

```
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB051009; BAB47561.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000197250; Homo sapiens.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     180    180
SQ   SEQUENCE   180 AA;  19710 MW;  4C8F196C9C788497 CRC64;

Alignment Scores:
Pred. No.:              24.4          Length:         180
Score:                  156.00        Matches:        156
Percent Similarity:     86.7%         Conservative:   0
Best Local Similarity:  86.7%         Mismatches:     24
Query Match:            6.4%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB2_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||   |||   ||||||||||||||||||
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerIleLeuIleGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                ||||||||||||||||||||||||||||||||||||||||||||||||||      |||
Db         21 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||||||||||||||||||||||      ||||||   ||||||||||||||||||
Db         41 ProAsnSerGlnCysIleArgTrpValSerHisProThrArgIleValCysLeuProSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrArgCysPheLysGlySerSerGluSer 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         81 IleCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                |||      |||||||||||||||||||||||||||||||   ||||||   |||
Db        101 AsnHisValValProLysProArgAsnLysArgValProIleLeuThrPheValValGly 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||   ||||||   ||||||||||||||||||||||||||||||||||||||||
Db        121 AlaGlyGlyLeuGlyGlyLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        141 TyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||   ||||||||||||   ||||||||||||||||
Db        161 ThrLeuGlnAspGlnLeuAsnSerLeuValAlaValValLeuGlnAsnArgArgAlaLeu 180

RESULT 18
Q96TB7_HUMAN
ID   Q96TB7_HUMAN   PRELIMINARY;   PRT;   180 AA.
AC   Q96TB7;
DT   01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT   01-DEC-2001, sequence version 1.
DT   07-FEB-2006, entry version 12.
DE   Envelope protein (Fragment).
GN   Name=env;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA   Kim H.-S., Lee W.-H.;
RT   "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT   identification, and phylogeny.";
RL   AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim H.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AB050999; BAB47556.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000198667; Homo sapiens.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002050; Env_polyprotein.
DR   Pfam; PF00429; TLV_coat; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     180    180
SQ   SEQUENCE   180 AA;  19984 MW;  3414BCDD74A060BA CRC64;

Alignment Scores:
Pred. No.:              25.5          Length:         180
Score:                  151.00        Matches:        151
Percent Similarity:     83.9%         Conservative:   0
Best Local Similarity:  83.9%         Mismatches:     29
Query Match:            6.2%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x Q96TB7_HUMAN (1-180)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerIleLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                ||||||||||||||||||||||||||||||||||||||||||||||||||      |||
Db         21 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||   ||||||||||||||||||||||||||||||   ||||||||||||||||
Db         41 ThrLysSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuProLeu 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||   |||      ||||||||||||||||||||   ||||||
Db         61 GlyIlePhePheValPheGlySerLeuAlaTyrArgCysLeuAsnGlyPheSerGluPro 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||||||||   ||||||||||||||   ||||||
Db         81 MetCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGlnTyrLeuTyr 100
```

```
Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        101 AsnTyrValIleProLysProHisAsnLysThrValLeuIleLeuProPheValIleArg 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        121 AlaGlyValLeuGlyGlyLeuSerThrGlyIleGlyGlyIleThrThrSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        141 TyrTyrLysLeuSerGlnGluLeuLysGlyAspMetGluGlnValThrAspSerLeuVal 160

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        161 ThrLeuGlnAspHisHisAsnSerLeuAlaAlaValValLeuGlnAsnGlnArgAlaLeu 180
```

RESULT 19
Q96TB3_HUMAN
ID Q96TB3_HUMAN PRELIMINARY; PRT; 179 AA.
AC Q96TB3;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Envelope protein (Fragment).
GN Name=env;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA Kim H.-S., Lee W.-H.;
RT "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT identification, and phylogeny.";
RL AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kim H.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC ---
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC ---
DR EMBL; AB051008; BAB47560.1; -; Genomic_DNA.
DR Ensembl; ENSG00000197330; Homo sapiens.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 179 179
SQ SEQUENCE 179 AA; 19471 MW; 70DF3E5658B8E667 CRC64;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 25.8 | Length: | 179 |
| Score: | 150.50 | Matches: | 161 |
| Percent Similarity: | 89.4% | Conservative: | 0 |
| Best Local Similarity: | 89.4% | Mismatches: | 18 |
| Query Match: | 6.2% | Indels: | 1 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x Q96TB3_HUMAN (1-179)

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---SerAsn 19

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
```

```
Db         20 LeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThr 39

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db         40 AlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuProSer 59

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db         60 GlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGluSer 79

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db         80 ValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeuTyr 99

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        100 AsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIleGly 119

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        120 AlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGlnPhe 139

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        140 TyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeuVal 159

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        160 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeu 179
```

RESULT 20
Q96TB8_HUMAN
ID Q96TB8_HUMAN PRELIMINARY; PRT; 179 AA.
AC Q96TB8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Envelope protein (Fragment).
GN Name=env;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21268768; PubMed=11375061; DOI=10.1089/088922201753342086;
RA Kim H.-S., Lee W.-H.;
RT "Human endogenous retrovirus HERV-W family: chromosomal localization,
RT identification, and phylogeny.";
RL AIDS Res. Hum. Retroviruses 17:643-648(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kim H.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC ---
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC ---
DR EMBL; AB050996; BAB47555.1; -; Genomic_DNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 179 179
SQ SEQUENCE 179 AA; 19228 MW; 3A9FACBAB6376E74 CRC64;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 26.5 | Length: | 179 |
| Score: | 147.50 | Matches: | 158 |

```
Percent Similarity:     87.8%          Conservative:  0
Best Local Similarity:  87.8%          Mismatches:    21
Query Match:            6.0%           Indels:        1
DB:                     1              Gaps:          1

US-10-717-580-11-COPY (1-7582) x Q96TB8_HUMAN (1-179)

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db          1 AsnAsnPheSerThrGluIleAsnThrThrSerIleLeuValGlyProLeuValSerAsn 20

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                ||||||   |||||||||||||||||||||||||||||||||||||      |||
Db         21 LeuGluIleIleHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAlaThr 40

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||||||||||||||||   ||||||||||||||   ||||||||||   |||
Db         41 AlaAsnSerGlnCysIleArgArgValThrProProThrArgIleValCysLeuThrSer 60

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSer 80

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||   ||||||||||||||   ||||||
Db         81 MetCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGlnAsnLeuTyr 100

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                |||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AsnHisValAlaProLysProArgAsnLysArgValProIleLeuProPheValIleGly 120

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||   ||||||||||||   ||||||||||||   ||||||||||||
Db        121 AlaGlyValLeuGlyGlyLeuGlyThrGlyProGlyGlyIleThrAlaSerThrGlnPhe 140

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ---TyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuVal 159

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||   |||      ||||||
Db        160 ThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnGlnThrAlaLeu 179


Search completed: August  4, 2006, 17:16:59
Job time : 27 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 4, 2006, 17:13:31 ; Search time 5 Seconds
(without alignments)
7.450 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 2447
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: UNITARY_AA
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 37 seqs, 16377 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 37 summaries

A-Geneseq

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=new.seq -DB=agsdb -SUFFIX=pto
-OUT=ags.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=unitary_AA -TRANS=human40.cdi -LIST=37 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=37 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : agsdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 678 | 27.7 | 846 | 1 | ABP41050 | HERV-7q env protei |
| 2 | 641 | 26.2 | 685 | 1 | ABP40988 | Human retroviral e |
| 3 | 518 | 21.2 | 540 | 1 | ABP40980 | Human retroviral H |
| 4 | 518 | 21.2 | 540 | 1 | AED74567 | Human placental pr |
| 5 | 517 | 21.1 | 538 | 1 | AAW99345 | Human endogenous r |
| 6 | 517 | 21.1 | 538 | 1 | AAB75138 | HERV-W envelope pr |
| 7 | 517 | 21.1 | 538 | 1 | AAB67652 | Amino acid sequenc |
| 8 | 517 | 21.1 | 538 | 1 | AED10943 | HERV envelope prot |
| 9 | 516 | 21.1 | 538 | 1 | AAY08622 | Human secreted pro |
| 10 | 516 | 21.1 | 538 | 1 | AAY67313 | Human secreted pro |
| 11 | 516 | 21.1 | 538 | 1 | AAE14540 | Human syncytin pro |
| 12 | 516 | 21.1 | 538 | 1 | ADC38777 | Human secreted pro |
| 13 | 516 | 21.1 | 538 | 1 | AED10985 | HERV R393E/Q envel |
| 14 | 515 | 21.0 | 538 | 1 | AED10987 | HERV R393E/Q F399A |
| 15 | 500.5 | 20.5 | 544 | 1 | ADE09243 | Novel protein-rela |
| 16 | 500.5 | 20.5 | 544 | 1 | ADF60170 | Human contig polyp |
| 17 | 448 | 18.3 | 542 | 1 | AAE25054 | Human EMBRY-2 prot |
| 18 | 441 | 18.0 | 542 | 1 | AAB19069 | Amino acid sequenc |
| 19 | 439 | 17.9 | 542 | 1 | AEB44408 | MSRV/HERV-W ENV-pV |
| 20 | 433.5 | 17.7 | 541 | 1 | AAB08195 | Amino acid sequenc |
| 21 | 416 | 17.0 | 493 | 1 | AAW71068 | Multiple sclerosis |
| 22 | 416 | 17.0 | 493 | 1 | AAW99553 | Protein encoded by |
| 23 | 416 | 17.0 | 493 | 1 | ADG14847 | MSRV associated pr |
| 24 | 411 | 16.8 | 493 | 1 | ADB84401 | MSRV-1 associated |
| 25 | 249 | 10.2 | 287 | 1 | AEB44407 | MSRV/HERV-W ENV-pS |
| 26 | 221.5 | 9.1 | 263 | 1 | AAM18008 | Peptide #4442 enco |
| 27 | 221.5 | 9.1 | 263 | 1 | ABB37043 | Peptide #4549 enco |
| 28 | 221.5 | 9.1 | 263 | 1 | AAM30519 | Peptide #4556 enco |
| 29 | 221.5 | 9.1 | 263 | 1 | ABB31812 | Peptide #4463 enco |
| 30 | 221.5 | 9.1 | 263 | 1 | ABB22358 | Protein #4357 enco |
| 31 | 221.5 | 9.1 | 263 | 1 | AAM70185 | Human bone marrow |
| 32 | 221.5 | 9.1 | 263 | 1 | AAM57770 | Human brain expres |
| 33 | 221.5 | 9.1 | 263 | 1 | ABG51885 | Human liver peptid |
| 34 | 221.5 | 9.1 | 263 | 1 | AAM05648 | Peptide #4330 enco |
| 35 | 221.5 | 9.1 | 263 | 1 | ABG39819 | Human peptide enco |
| 36 | 111 | 4.5 | 120 | 1 | AAB75137 | HERV-W truncated e |
| 37 | 102 | 4.2 | 122 | 1 | ADK35467 | Novel human polype |

ALIGNMENTS

RESULT 1
ABP41050
ID ABP41050 standard; protein; 846 AA.
XX
AC ABP41050;
XX
DT 01-AUG-2002 (first entry)
XX
DE HERV-7q env protein #1.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis.
XX
OS Human endogenous retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Example 4; Fig 19; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 846 AA;

Alignment Scores:
Pred. No.: 0.059 Length: 846

Score: 678.00 Matches: 808
Percent Similarity: 93.8% Conservative: 0
Best Local Similarity: 93.8% Mismatches: 38
Query Match: 27.7% Indels: 15
DB: 1 Gaps: 11

US-10-717-580-11-COPY (1-7582) x ABP41050 (1-846)

```
Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTT 4730
                ||||||||||||||||||||   ||||||||||||||   ||||||||||||   |||
Db          1 LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn---Phe 19

Qy       4731 TACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCAC 4790
                ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         20 TyrProGlnAlaGlnThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHis 39

Qy       4791 GGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACT 4850
                |||||||||   ||||||||||||   |||||||||||||||||||||||||||||||||
Db         40 GlyLeuGlyArgGlyLeuProLeu---AspArgLysGlyProArgGlyAsnLysGlyThr 58

Qy       4851 AGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCC 4910
                ||||||   |||||||||||||||||||||||||||||||||||||||||||   |||
Db         59 SerSer---AsnAsnSerGlnIleArgThrSerProArgLeuThrGlu------GlnPro 75

Qy       4911 TGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTT 4970
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 CysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu 95

Qy       4971 ACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAA 5030
                ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         96 ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu---AsnThrGln 114

Qy       5031 AGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTAT 5090
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 ArgThrSerLysLysAlaAsnProGlyAsnProProHisMetAlaCysSerValAlaTyr 134

Qy       5091 AGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCT 5150
                ||||||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        135 SerLeuLysLysAsnLeuGlnLeuSerProLysSerArgThr---ProIleArgAsnAla 153

Qy       5151 GTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTT 5210
                ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        154 ValTrpLysAlaLeuHisAsnGln------ProCysAlaProLysThrAlaAsnLeuVal 171

Qy       5211 GCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTAT 5270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 AlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr 191

Qy       5271 CCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTT 5330
                |||   ||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        192 Pro---GluGluGlyLysGluLeuPheHisProCysAspMetValLeuValLysSerLeu 210

Qy       5331 CYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYT 5390
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ProSerAsnSerProSerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSer 230

Qy       5391 ACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAA 5450
                ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        231 ThrProThrAlaValLysValAlaGlyValGluSerTrpIleHisHisThr---ValLys 249

Qy       5451 TCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAA 5510
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 SerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAlaSerTyrSerCysGlu 269

Qy       5511 CCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCA 5570
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 ProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys------LeuLys 287

Qy       5571 TAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTT 5630
                   ||| ||||||   ||||||||||||||||||||     |||     |||   ||
Db        288 SerIle-ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPh 307

Qy       5631 CACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTT 5690

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        307 eThrLeuThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPh 327

Qy       5691 TCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAA 5750
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        327 eLeuTrpArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLy 347

Qy       5751 GGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCAC 5810
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 sGlyThrProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaTh 367

Qy       5811 TCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCC 5870
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 rLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysPr 387

Qy       5871 TGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGG 5930
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 oGlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGl 407

Qy       5931 GGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCAC 5990
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 yGlyGlyValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuTh 427

Qy       5991 CSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 6050
                |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 rArgValHisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluTh 447

Qy       6051 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGA 6110
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 rLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGl 467

Qy       6111 GGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATA 6170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        467 uValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTy 487

Qy       6171 TGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTC 6230
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 rValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSe 507

Qy       6231 CGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTG 6290
                ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        507 rValLeuValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCy 527

Qy       6291 TGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCC 6350
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 sValLysPheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrPr 547

Qy       6351 TCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTA 6410
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 oProThrGlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTy 567

Qy       6411 TCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYAT 6470
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 rArgCysLeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMe 587

Qy       6471 GRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAG 6530
                |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 tThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysAr 607

Qy       6531 AGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAT 6590
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 gValProIleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIl 627

Qy       6591 TGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGA 6650
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 eGlyGlyIleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAs 647

Qy       6651 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 6710
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 pMetGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAl 667

Qy       6711 AGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTG 6770
                ||||||||||   ||||||||||||||||   ||||||||||||||||||||||||||||
```

```
Db        667 aValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCy 687

Qy       6771 TTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAA 6830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 sLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLy 707

Qy       6831 AGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACC 6890
             ||||  ||||||  ||||||||||||||  ||||||||||||||||||||||||||||
Db        707 sValLysGluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyPr 727

Qy       6891 CTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGC 6950
             ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        727 oTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAl 747

Qy       6951 TATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTC 7010
             ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        747 aIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSe 767

Qy       7011 TTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGAT 7070
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 rSerArgIleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIl 787

Qy       7071 CTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGG 7130
             ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        787 eTyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGl 807

Qy       7131 CACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAG 7190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 yThrProProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySe 827

Qy       7191 CAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGG 7250
             ||||  ||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        827 rSer---SerGlyArgArgProThrSerProThrAlaLeuArgPheSerCysAspGlyGl 846

Qy       7251 G 7251
             |
Db        846 y 846
```

RESULT 2
ABP40988
ID ABP40988 standard; protein; 685 AA.
XX
AC ABP40988;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral env ORF protein.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis.
XX
OS Human endogenous retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Claim 22; Page 152-154; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
SQ Sequence 685 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.0941 | Length: | 685 |
| Score: | 641.00 | Matches: | 654 |
| Percent Similarity: | 95.5% | Conservative: | 0 |
| Best Local Similarity: | 95.5% | Mismatches: | 30 |
| Query Match: | 26.2% | Indels: | 1 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x ABP40988 (1-685)

```
Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ProLysThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeu 20

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
             ||||||||||||||||||||||||  ||||||||||||||||||||||||||||  |||
Db         21 LysThrLeuGlnGlyThrTyrPro***GluGluGlyLysGluLeuPheHisProCysAsp 40

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db         41 MetValLeuValLysSerLeuProSerAsnSerProSerLeuAspThrSerTrpGluGly 60

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db         61 ProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyValGluSerTrp 80

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
             ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleHisHisThr***ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAsp 100

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AsnAlaSerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnPro 120

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
             ||||||||  ||||||||  ||| ||||||  ||||||||||||||||||||||||||||
Db        121 GlyGlyLys***LeuLysSer***Ile-ProMetAlaLeuProTyrHisIlePheLeuPh 140

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
                 |||   |||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        140 eThrValLeuLeuProSerPheThrLeuThrAlaProProProCysArgCysMetThrSe 160

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
             ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 rSerSerProTyrGlnGluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaPr 180

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
             ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        180 oSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMetProAr 200

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 gAsnCysTyrHisSerAlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLy 220

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 sMetIleAsnProSerCysProGlyGlyLeuGlyValThrValCysTrpThrTyrPheTh 240
```

```
Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 rGlnThrGlyMetSerAspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLy 260

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        260 sGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLysGlyLeuAs 280

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 pLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnTh 300

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 rThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLe 320

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
             ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        320 uProLeuAsnPheArgProTyrValSerIleProValProGluGlnTrpAsnAsnPheSe 340

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
             ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        340 rThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsnLeuGluIleTh 360

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 rHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGl 380

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 nCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePh 400

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 eValCysGlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLe 420

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        420 uSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIl 440

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 eSerLysProArgAsnLysArgValProIleLeuProPheValIleGlyAlaGlyValLe 460

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 uGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyrTyrLysLe 480

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 uSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuValThrLeuGlnAs 500

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
             ||||||||||||||||||||||||||||||   ||||||||||||||||||   |||||
Db        500 pGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuTh 520

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 rAlaGluArgGlyGlyThrCysLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGl 540

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
             ||||||||||||||||||||||||   ||||||   ||||||||||||||   |||||||
Db        540 nSerGlyIleValThrGluLysValLysGluIleArgAspArgIleGlnArgArgAlaGl 560

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
             ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        560 uGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuPr 580

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 oPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuPheGlyProCysIlePheAs 600
```

```
Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 nLeuLeuValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLy 620

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        620 sMetGlnSerLysThrLysIleTyrArgArgProLeuAspArgProAlaSerProArgSe 640

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 rAspValAsnAspIleLysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLe 660

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             ||||||||||||||||||||||||   ||||||   ||||||||||||||||||||||||
Db        660 uArgProAsnSerAlaGlySerSer***SerGlyArgArgProThrSerProThrAlaLe 680

Qy       7230 TAGGTTTTCCTGT 7242
             |||||||||||||
Db        680 uArgPheSerCys 684
```

RESULT 3
ABP40980
ID ABP40980 standard; protein; 540 AA.
XX
AC ABP40980;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral HERV-7q env peptide #3.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis.
XX
OS Human endogenous retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Claim 22; Fig 4; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
SQ Sequence 540 AA;

Alignment Scores:
Pred. No.: 0.45 Length: 540

| | | | |
|---|---|---|---|
| Score: | 518.00 | Matches: | 518 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.2% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x ABP40980 (1-540)

```
Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540
```

RESULT 4
AED74567
ID AED74567 standard; protein; 540 AA.
XX
AC AED74567;
XX
DT 12-JAN-2006 (first entry)
XX
DE Human placental protein SEQ ID NO:1395.
XX
KW pre-eclampsia; gynecological.
XX
OS Homo sapiens.
XX
PN US2005255114-A1.
XX
PD 17-NOV-2005.
XX
PF 07-APR-2004; 2004US-00821234.
XX
PR 07-APR-2003; 2003US-0462047P.
XX
PA (NUVE-) NUVELO INC.
XX
PI Labat I, Tang YT, Stache-Crain B, Boyle B;
XX
DR WPI; 2005-808574/82.
DR N-PSDB; AED73715.
XX
PT Identifying a patient with a higher risk of preeclampsia comprises

PT nucleic acid hybridization assay or antibody assay to determine level of
PT a specific nucleic acid (mRNA) or polypeptide.
XX
PS Claim 2; SEQ ID NO 1395; 358pp; English.
XX
CC The invention relates to a method for identifying a patient with a higher
CC risk of preeclampsia. The method comprises: (a) assaying a sample from
CC the patient to determine the level of a mRNA or other polynucleotide
CC within the sample that hybridizes specifically to a polynucleotide of
CC AED73173-AED74024, and comparing the level to a standard; or (b)
CC contacting a sample from the patient with an antibody that specifically
CC binds to a polypeptide of AED74025-AED74876 to determine the level of
CC polypeptide within the sample, and comparing the level to a standard.
CC Also described: (1) a diagnostic kit, for detecting preeclampsia,
CC comprising: (a) an antibody specific for any of the polypeptides of
CC AED74025-AED74876 or their fragments; or (b) a polynucleotide sequence
CC comprising any of AED73173-AED74024 coupled to a surface; and (c) a
CC standard for any of the polypeptides of AED74025-AED74876 or any of the
CC polynucleotides of AED73173-AED74024 indicative of a higher risk of
CC diagnosis of preeclampsia; (2) a pharmaceutical composition comprising an
CC antibody specific for any of the polypeptides of AED74025-AED74876 or one
CC or more purified polypeptides of AED74025-AED74876, to ameliorate signs
CC or symptoms of preeclampsia; and (3) a method of treating preeclampsia.
CC The method and kit are useful for identifying a patient with a higher
CC risk of preeclampsia. The composition and method are useful for treating
CC preeclampsia. The present sequence represents an amino acid sequence
CC encodes by a nucleic acid sequence obtained from a human placental cDNA
CC library. Note - The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from the USPTO web site.
XX
SQ Sequence 540 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.45 | Length: | 540 |
| Score: | 518.00 | Matches: | 518 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.2% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x AED74567 (1-540)

```
Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
```

```
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521
```

```
Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540

RESULT 5
AAW99345
ID   AAW99345 standard; protein; 538 AA.
XX
AC   AAW99345;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W ORF1.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Example 8; Fig 6; 106pp; French.
XX
CC   This sequence represents the open reading frame 1 encoding the env
CC   protein from the human endogenous retrovirus (HERV) W (AAX25665). Nucleic
CC   acids, their fragments or peptides encoded by them derived from the HERV-
CC   W genome are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 538 AA;

Alignment Scores:
Pred. No.:              0.456          Length:         538
Score:                  517.00         Matches:        517
Percent Similarity:     96.1%          Conservative:   0
Best Local Similarity:  96.1%          Mismatches:     21
Query Match:            21.1%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x AAW99345 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||||||||||||      |||     |||   ||||||||||||
Db          1 MetGlyLeuProTyrHisIlePheLeuCysSerValLeuSerProCysPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProHisProGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlyValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnValGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        381 ArgAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValGluGlu 420
```

```
Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
                |||   |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        421 IleProAspArgIleGlnArgIleAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
                ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerArgTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
                ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAspLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
                |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 6
AAB75138
ID AAB75138 standard; protein; 538 AA.
XX
AC AAB75138;
XX
DT 08-AUG-2001 (first entry)
XX
DE HERV-W envelope protein G.
XX
KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW envelope protein; multiple sclerosis-related superantigen; vaccine;
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW antisense-therapy; autoimmune disorder.
XX
OS Human endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT Region 271. .317
FT /label= SU
FT /note= "surface protein"
FT Cleavage-site 317. .318
FT /note= "SU-TM putative cleavage site"
FT Region 318. .538
FT /label= TM
FT /note= "transmembrane subunit"
XX
PN WO200131021-A1.
XX
PD 03-MAY-2001.
XX
PF 30-OCT-2000; 2000WO-EP010659.
XX
PR 28-OCT-1999; 99EP-00402690.
XX
PA (UYGE-) UNIV GENEVE.
XX
PI Conrad B, Mach B;
XX
DR WPI; 2001-316336/33.
DR N-PSDB; AAH20070.
XX
PT New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT activity useful for diagnosing and treating multiple sclerosis.
XX
PS Claim 1; Fig 7; 94pp; English.
XX
CC On the basis of the PBS t-RNA motif used for the classification of human
CC endogenous retrovirus (HERVs) the full length endogenous provirus which
CC was been located on the long arm of human chromosome 7 (7q21-22) has been
CC designated HERV-W. The present invention describes proteins or peptides
CC (I) having superantigen (SAg) activity comprising the ENV protein (ENV)
CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC have neuroprotective activity, and can be used in: vaccines; antisense-
CC therapy; and HERV-W SAg activity-inhibitors. (I) and encoding DNA/RNA are
CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC disorders. (I) are also useful for identifying substances (and optionally
CC recovering) capable of binding to a retroviral superantigen associated
CC with MS, substances capable of blocking SAg activity and substances
CC capable of blocking transcription or translation of HERV-W retroviral
CC superantigen. A protein or peptide derived from (I), modified to be
CC devoid of SAg activity and being capable of generating an immune response
CC against HERV-W retroviral SAg is useful in therapy. Nucleic acid
CC molecules encoding (I) are useful as vaccines against MS. Substances
CC capable of blocking SAg activity, capable of binding to a retroviral
CC superantigen associated with MS, or capable of blocking transcription or
CC translation of HERV-W retroviral superantigen for use in treating or
CC preventing MS, obtained using (I) are useful for the treatment and
CC prevention of MS. (I) and nucleic acids encoding them are useful for
CC diagnosing autoimmune disease. The present sequence represents the
CC specifically claimed envelope protein of HERV-W designated G
XX
SQ Sequence 538 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.456 | Length: | 538 |
| Score: | 517.00 | Matches: | 517 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x AAB75138 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
                |||   |||||||||||||||||||      |||      |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
                |||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
                |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
                ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
```

```
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240

Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300

Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360

Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420

Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480

Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540

Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600

Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660

Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720

Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780

Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840

Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960

Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020

Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080

Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140

Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194

Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 7
AAB67652
```

```
ID   AAB67652 standard; protein; 538 AA.
XX
AC   AAB67652;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a human endogenous retrovirus envelope protein.
XX
KW   Envelope protein; HERV; syncytia formation; placental development;
KW   syncytia; cancer; cell adhesion.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200116171-A1.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-FR002429.
XX
PR   01-SEP-1999;   99FR-00011141.
PR   15-SEP-1999;   99FR-00011793.
XX
PA   (INMR ) BIO MERIEUX.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Mallet F,  Cosset F,  Blond J,  Lavillette D,  Bouton O,  Ruggieri A;
XX
DR   WPI; 2001-226676/23.
DR   N-PSDB; AAF55630.
XX
PT   Detecting expression of human endogenous retrovirus envelope protein in
PT   cells of a tissue or culture, from its ability to induce syncytia.
XX
PS   Claim 1; Page 41-44; 57pp; French.
XX
CC   The present sequence represents a human endogenous retrovirus envelope
CC   protein. The specification describes a method for detecting expression of
CC   an envelope protein from a human endogenous retrovirus (HERV), in cells,
CC   of a tissue or culture. The method comprises detecting syncytia formation
CC   due to the fusogenic properties of the envelope protein. Envelope
CC   polypeptides and polynucleotides are used to produce therapeutic or
CC   prophylactic compositions, particularly for treatment of cancer, to
CC   correct defects in placental development (or other natural formation of
CC   other types of syncytia), and to promote adhesion of cells in grafts or
CC   cellular repair processes. Expression of sequences antisense to the
CC   polynucleotide are used to prevent formation of syncytia
XX
SQ   Sequence 538 AA;
```

```
Alignment Scores:
Pred. No.:                0.456          Length:          538
Score:                    517.00         Matches:         517
Percent Similarity:       96.1%          Conservative:    0
Best Local Similarity:    96.1%          Mismatches:      21
Query Match:              21.1%          Indels:          0
DB:                       1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x AAB67652 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640

Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700

Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760

Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
```

```
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                 ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 8
AED10943
ID AED10943 standard; protein; 538 AA.
XX
AC AED10943;
XX
DT 15-DEC-2005 (first entry)
XX
DE HERV envelope protein SEQ ID NO 74.
XX
KW virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW screening; immune modulation; graft rejection; allergy;
KW autoimmune disease; cancer; viral infection; envelope; env.
XX
OS Human endogenous retrovirus W.
XX
PN WO2005095442-A1.
XX
PD 13-OCT-2005.
XX
PF 30-MAR-2005; 2005WO-EP003339.
XX
PR 30-MAR-2004; 2004EP-00290838.
XX
PA (INSR ) INST ROUSSY GUSTAVE.
PA (CNRS ) CENT NAT RECH SCI.
PA (UYPA-) UNIV PARIS SUD.
XX
PI Renard M, Mangeney M, Heidmann T;
XX
DR WPI; 2005-703311/72.
DR N-PSDB; AED10942.
XX
PT New polypeptide capable of modulating the immunosuppressive properties of
PT a viral protein or its fragment, useful for the manufacture of a vaccine
PT for treating or preventing viral infections or cancer.
XX
PS Disclosure; SEQ ID NO 74; 262pp; English.
XX
CC The invention describes a polypeptide of 7-20 amino acid residues, which
CC is capable of modulating the immunosuppressive properties of a viral
CC protein or its fragment, against the host in which it is expressed
CC (immunosuppression-modulatory sequence) when it substitutes the
CC homologous sequence of the viral protein or its fragment, the polypeptide
CC comprising the minimum following consensus amino acid sequence of (I).
CC Also described are: a mutated ENV protein resulting from the mutation of
CC a wild type ENV protein essentially carrying the following sequence: NY 1
CC Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC or prokaryotic expression vector comprising the nucleic acid as well as
CC the elements necessary for the expression of the nucleic acid; a
CC recombinant cell comprising the nucleic acid or the eukaryotic or

CC prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC comprising as active substance: at least one polypeptide, mutated ENV
CC protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC expression vector, or recombinant cell; in association with a
CC pharmaceutically acceptable carrier; and antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides, directed against
CC mutated ENV proteins, provided that the antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides do not bind to the
CC corresponding wild type ENV proteins. The first mutation of a first amino
CC acid and optionally of a second mutation of a second amino acid in a wild
CC type viral envelope (ENV) protein is useful for manufacturing a mutated
CC ENV protein having a modified immunosuppressive activity with respect to
CC the wild type ENV protein. The protein is useful for the manufacture of a
CC medicament or a vaccine intended for the prevention and/or the treatment
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC requiring an inhibition of the immune system, such as autoimmune
CC diseases, allergies or graft rejections. The protein is also useful for
CC screening compounds liable to modulate the immunosuppressive activity of
CC viruses or tumor cells, where the compounds to screen are peptides, in
CC particular peptides comprising from 5-30 amino acids, such as peptides
CC originating from combinatorial peptide libraries. This is the amino acid
CC sequence of HERV envelope protein.
XX
SQ Sequence 538 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.456 | Length: | 538 |
| Score: | 517.00 | Matches: | 517 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x AED10943 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 9
AAY08622
ID AAY08622 standard; protein; 538 AA.

```
XX
AC   AAY08622;
XX
DT   10-AUG-1999  (first entry)
XX
DE   Human secreted protein AJ172_2.
XX
KW   Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW   bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW   cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW   haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW   cadherin; tumour invasion suppressor; gene therapy; tissue growth.
XX
OS   Homo sapiens.
XX
PN   WO9926972-A1.
XX
PD   03-JUN-1999.
XX
PF   17-NOV-1998;   98WO-US024614.
XX
PR   21-NOV-1997;   97US-00976110.
PR   18-MAY-1998;   98US-00080478.
PR   20-OCT-1998;   98US-00175928.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR   WPI; 1999-357813/30.
DR   N-PSDB; AAX77526.
XX
PT   New polynucleotides encoding secreted proteins.
XX
PS   Claim 14a; Page 101-103; 142pp; English.
XX
CC   This invention describes novel human secreted proteins encoded by
CC   polynucleotides isolated from human adult testes, adult brain, adult
CC   blood or adult placenta, or murine adult bone marrow or thymus cDNA
CC   libraries. The products of the invention are predicted to have biological
CC   activities which would make them suitable for treating, preventing or
CC   ameliorating medical conditions in humans and animals, although no
CC   supporting data is given. Suggested activities include nutritional
CC   activity, cytokine and cell proliferation/differentiation activity,
CC   immune stimulating (e.g. as vaccines) or suppressing activity,
CC   haematopoiesis regulating activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, cadherin/tumour invasion suppressor activity, and tumour
CC   inhibition activity. The polynucleotides are also stated to be useful for
CC   gene therapy
XX
SQ   Sequence 538 AA;

Alignment Scores:
Pred. No.:              0.462          Length:          538
Score:                  516.00         Matches:         516
Percent Similarity:     95.9%          Conservative:    0
Best Local Similarity:  95.9%          Mismatches:      22
Query Match:            21.1%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x AAY08622 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||||||||||||      |||      |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400
```

```
Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             |||   |||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 10
AAY67313
ID   AAY67313 standard; protein; 538 AA.
XX
AC   AAY67313;
XX
DT   11-APR-2000  (first entry)
XX
DE   Human secreted protein AJ172_2 amino acid sequence.
XX
KW   Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW   placental pathology; metastasis inhibition; nutritional activity;
KW   immune stimulator; haematopoiesis regulator; tissue growth;
KW   tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115.
XX
OS   Homo sapiens.
XX
PN   WO9960020-A1.
XX
PD   25-NOV-1999.
XX
PF   17-MAY-1999;   99WO-US010915.
XX
PR   18-MAY-1998;   98US-00080478.
PR   20-OCT-1998;   98US-00175928.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Mi S,  Treacy M;
XX
DR   WPI; 2000-116311/10.
DR   N-PSDB; AAZ59468.
XX
PT   New polynucleotides encoding secreted cDNA libraries, used to develop
PT   products for the diagnosis and treatment of neoplastic disease.
XX
PS   Claim 15; Page 108-110; 149pp; English.
XX
CC   This is the human secreted protein AJ172_2 amino acid sequence, the
CC   polynucleotide encoding AJ172_2 was obtained from a human adult testes
CC   cDNA library. The invention relates to secreted human and murine
CC   proteins. The polynucleotides and proteins are predicted to have
CC   biological activities which would make them suitable for treating,
CC   preventing or ameliorating medical conditions in humans and animals.
CC   Detection of the levels of the proteins can be used for the diagnosis of
CC   e.g. pre-eclampsia, placental pathology or cancer. Agents which modulate
CC   the expression or function of the proteins may be used for treating a
CC   neoplastic disease and inhibiting metastasis. Other suggested activities
CC   include nutritional activity (e.g. in feeds), cytokine and cell
CC   proliferation/differentiation activity, immune stimulating (e.g. as
CC   vaccines) or suppressing activity, haematopoiesis regulating activity,
CC   tissue growth activity, activin/inhibin activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC   receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC   invasion suppressor activity, and tumour inhibition activity. The
CC   polynucleotide sequences are also stated to be useful for gene therapy
XX
SQ   Sequence 538 AA;

Alignment Scores:
Pred. No.:                0.462          Length:          538
Score:                    516.00         Matches:         516
Percent Similarity:       95.9%          Conservative:    0
Best Local Similarity:    95.9%          Mismatches:      22
Query Match:              21.1%          Indels:          0
DB:                       1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x AAY67313 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||  ||||||||||||||||||||      |||     |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240
```

```
Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 11
AAE14540
ID AAE14540 standard; protein; 538 AA.
XX
AC AAE14540;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human syncytin protein.
XX
KW Human; syncytin; preeclampsia; gestational trophoblast disorder;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W.
XX
OS Homo sapiens.
XX
PN WO200204678-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US021719.
XX
PR 07-JUL-2000; 2000US-0216657P.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Keith JC, Mccoy JM, Mi S;
XX
DR WPI; 2002-171727/22.
DR N-PSDB; AAD24195.
XX
PT Identifying a compound for treating a subject with or at risk of
PT developing preeclampsia, comprises determining whether the expression or
PT activity of syncytin in the cell is modulated in the presence of a test
PT compound.
XX
PS Disclosure; Page 42-43; 43pp; English.
XX
CC The invention relates to identifying compounds which are modulators of
CC syncytin expression. The syncytin modulators are useful in diagnosis and
CC treatment of preeclampsia and gestational trophoblast disorders (e.g.
CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the envelope
CC gene of human endogenous defective retrovirus, HERV-W. The present
CC invention is based partly on the discovery that syncytin expression is
CC dramatically reduced in preeclampsia, and is also mis-localised to the
CC apical syncytiotrophoblast membrane. The present sequence is human
CC syncytin protein
XX
SQ Sequence 538 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.462 | Length: | 538 |
| Score: | 516.00 | Matches: | 516 |
| Percent Similarity: | 95.9% | Conservative: | 0 |
| Best Local Similarity: | 95.9% | Mismatches: | 22 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x AAE14540 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
```

```
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060

Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120

Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180

Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240

Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300

Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360

Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420

Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480

Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540

Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600

Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660

Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720

Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780

Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840

Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960

Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020

Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080

Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140

Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194

Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 12
ADC38777
ID  ADC38777 standard; protein; 538 AA.
XX
AC  ADC38777;
XX
DT  18-DEC-2003  (first entry)
XX
DE  Human secreted protein #63.
XX
KW  immune disorder; severe combined immunodeficiency; SCID;
KW  autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW  rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
KW  lymphoid cell deficiency; osteoporosis; osteoarthritis;
KW  peripheral nervous system disease; peripheral neuropathy;
KW  Alzheimer's disease; Parkinson's disease; coagulation disorder;
KW  inflammatory disease; systemic inflammatory response syndrome; SIRS;
KW  ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;
KW  hypersensitivity; regeneration; neural cell proliferation; fertility;
KW  tumour; chemokine; human; secreted protein.
XX
OS  Homo sapiens.
XX
PN  US2002193567-A1.
XX
PD  19-DEC-2002.
XX
PF  02-APR-2002; 2002US-00114893.
XX
PR  11-AUG-1995;  95US-00514014.
PR  05-APR-1996;  96US-00628364.
PR  19-APR-1996;  96US-00635311.
PR  07-JUN-1996;  96US-00659224.
PR  17-JUN-1996;  96US-00664596.
PR  09-JUL-1996;  96US-00677231.
PR  26-JUL-1996;  96US-00686878.
PR  23-AUG-1996;  96US-00701819.
PR  27-SEP-1996;  96US-00721488.
PR  27-SEP-1996;  96US-00721798.
PR  27-SEP-1996;  96US-00721923.
PR  27-SEP-1996;  96US-00721926.
PR  25-OCT-1996;  96US-00738367.
PR  30-OCT-1996;  96US-00739775.
PR  13-JAN-1997;  97US-00783395.
PR  10-APR-1997;  97US-00833823.
PR  02-JUN-1997;  97US-00867677.
PR  05-SEP-1997;  97US-00924838.
PR  06-OCT-1999;  99US-00413232.
XX
PA  (GEMY ) GENETICS INST INC.
XX
PI  Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI  Merberg D,  Treacy M,  Bowman MR,  Spaulding V,  Carlin-Duckett M;
PI  Kelleher K;
XX
DR  WPI; 2003-657236/62.
DR  N-PSDB; ADC38776.
XX
PT  Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
PT  BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
PT  useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
XX
PS  Disclosure; SEQ ID NO 135; 412pp; English.

XX
CC The invention relates to a protein comprising fully defined AZ302 1
CC protein or BD127 1 6 protein. The polynucleotides are useful for
CC expressing recombinant proteins for analysis and are also useful as
CC chromosome markers or tags to identify chromosomes or to map related gene
CC positions. The proteins are useful as amino acid supplement, carbon
CC source, nitrogen source and carbohydrate source. The proteins are useful
CC for treating various immune deficiencies and disorders (e.g. severe
CC combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),
CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
CC nerve growth or regeneration, for proliferating neural cells and for
CC regenerating nerve and brain tissue, for inducing fertility and for
CC inhibiting tumour growth. Proteins are also useful as chemokine for
CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
CC useful as inhibitors of receptor/ligand interactions. The present
CC sequence represents the amino acid sequence of a human secreted protein.
XX
SQ Sequence 538 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.462 | Length: | 538 |
| Score: | 516.00 | Matches: | 516 |
| Percent Similarity: | 95.9% | Conservative: | 0 |
| Best Local Similarity: | 95.9% | Mismatches: | 22 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x ADC38777 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   ||||||||||||||||||||      |||      |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 13
AED10985

ID AED10985 standard; protein; 538 AA.
XX
AC AED10985;
XX
DT 15-DEC-2005 (first entry)
XX
DE HERV R393E/Q envelope protein SEQ ID NO 116.
XX
KW virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW screening; immune modulation; graft rejection; allergy;
KW autoimmune disease; cancer; viral infection; envelope; env; mutein.
XX
OS Human endogenous retrovirus W.
XX
PN WO2005095442-A1.
XX
PD 13-OCT-2005.
XX
PF 30-MAR-2005; 2005WO-EP003339.
XX
PR 30-MAR-2004; 2004EP-00290838.
XX
PA (INSR ) INST ROUSSY GUSTAVE.
PA (CNRS ) CENT NAT RECH SCI.
PA (UYPA-) UNIV PARIS SUD.
XX
PI Renard M, Mangeney M, Heidmann T;
XX
DR WPI; 2005-703311/72.
DR N-PSDB; AED10984.
XX
PT New polypeptide capable of modulating the immunosuppressive properties of
PT a viral protein or its fragment, useful for the manufacture of a vaccine
PT for treating or preventing viral infections or cancer.
XX
PS Claim 38; SEQ ID NO 116; 262pp; English.
XX
CC The invention describes a polypeptide of 7-20 amino acid residues, which
CC is capable of modulating the immunosuppressive properties of a viral
CC protein or its fragment, against the host in which it is expressed
CC (immunosuppression-modulatory sequence) when it substitutes the
CC homologous sequence of the viral protein or its fragment, the polypeptide
CC comprising the minimum following consensus amino acid sequence of (I).
CC Also described are: a mutated ENV protein resulting from the mutation of
CC a wild type ENV protein essentially carrying the following sequence: NY 1
CC Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC or prokaryotic expression vector comprising the nucleic acid as well as
CC the elements necessary for the expression of the nucleic acid; a
CC recombinant cell comprising the nucleic acid or the eukaryotic or
CC prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC comprising as active substance: at least one polypeptide, mutated ENV
CC protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC expression vector, or recombinant cell; in association with a
CC pharmaceutically acceptable carrier; and antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides, directed against
CC mutated ENV proteins, provided that the antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides do not bind to the
CC corresponding wild type ENV proteins. The first mutation of a first amino
CC acid and optionally of a second mutation of a second amino acid in a wild
CC type viral envelope (ENV) protein is useful for manufacturing a mutated
CC ENV protein having a modified immunosuppressive activity with respect to
CC the wild type ENV protein. The protein is useful for the manufacture of a
CC medicament or a vaccine intended for the prevention and/or the treatment
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC requiring an inhibition of the immune system, such as autoimmune
CC diseases, allergies or graft rejections. The protein is also useful for
CC screening compounds liable to modulate the immunosuppressive activity of
CC viruses or tumor cells, where the compounds to screen are peptides, in
CC particular peptides comprising from 5-30 amino acids, such as peptides
CC originating from combinatorial peptide libraries. This is the amino acid
CC sequence of HERV envelope protein.
XX

SQ Sequence 538 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.462 | Length: | 538 |
| Score: | 516.00 | Matches: | 516 |
| Percent Similarity: | 95.9% | Conservative: | 0 |
| Best Local Similarity: | 95.9% | Mismatches: | 22 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x AED10985 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
                ||| |||||||||||||||||||||      |||      |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
                |||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
                |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
                |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
                |||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
                |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
                |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
```

```
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGlu***GlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

```
RESULT 14
AED10987
ID   AED10987 standard; protein; 538 AA.
XX
AC   AED10987;
XX
DT   15-DEC-2005  (first entry)
XX
DE   HERV R393E/Q F399A envelope protein SEQ ID NO 118.
XX
KW   virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW   screening; immune modulation; graft rejection; allergy;
KW   autoimmune disease; cancer; viral infection; envelope; env; mutein.
XX
OS   Human endogenous retrovirus W.
XX
PN   WO2005095442-A1.
XX
PD   13-OCT-2005.
XX
PF   30-MAR-2005; 2005WO-EP003339.
XX
PR   30-MAR-2004; 2004EP-00290838.
XX
PA   (INSR ) INST ROUSSY GUSTAVE.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (UYPA-) UNIV PARIS SUD.
XX
PI   Renard M,  Mangeney M,  Heidmann T;
XX
DR   WPI; 2005-703311/72.
DR   N-PSDB; AED10986.
XX
PT   New polypeptide capable of modulating the immunosuppressive properties of
PT   a viral protein or its fragment, useful for the manufacture of a vaccine
PT   for treating or preventing viral infections or cancer.
XX
PS   Claim 38; SEQ ID NO 118; 262pp; English.
XX
CC   The invention describes a polypeptide of 7-20 amino acid residues, which
CC   is capable of modulating the immunosuppressive properties of a viral
CC   protein or its fragment, against the host in which it is expressed
CC   (immunosuppression-modulatory sequence) when it substitutes the
CC   homologous sequence of the viral protein or its fragment, the polypeptide
CC   comprising the minimum following consensus amino acid sequence of (I).
CC   Also described are: a mutated ENV protein resulting from the mutation of
CC   a wild type ENV protein essentially carrying the following sequence: NY 1
CC   Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC   acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC   or prokaryotic expression vector comprising the nucleic acid as well as
CC   the elements necessary for the expression of the nucleic acid; a
CC   recombinant cell comprising the nucleic acid or the eukaryotic or
CC   prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC   comprising as active substance: at least one polypeptide, mutated ENV
CC   protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC   expression vector, or recombinant cell; in association with a
CC   pharmaceutically acceptable carrier; and antibodies or their fragments,
CC   scFv polypeptides, aptamers, or binding peptides, directed against
CC   mutated ENV proteins, provided that the antibodies or their fragments,
CC   scFv polypeptides, aptamers, or binding peptides do not bind to the
CC   corresponding wild type ENV proteins. The first mutation of a first amino
CC   acid and optionally of a second mutation of a second amino acid in a wild
CC   type viral envelope (ENV) protein is useful for manufacturing a mutated
CC   ENV protein having a modified immunosuppressive activity with respect to
CC   the wild type ENV protein. The protein is useful for the manufacture of a
CC   medicament or a vaccine intended for the prevention and/or the treatment
CC   of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC   requiring an inhibition of the immune system, such as autoimmune
CC   diseases, allergies or graft rejections. The protein is also useful for
CC   screening compounds liable to modulate the immunosuppressive activity of
CC   viruses or tumor cells, where the compounds to screen are peptides, in
CC   particular peptides comprising from 5-30 amino acids, such as peptides
CC   originating from combinatorial peptide libraries. This is the amino acid
CC   sequence of HERV envelope protein.
XX
SQ   Sequence 538 AA;
```

```
Alignment Scores:
Pred. No.:              0.468          Length:          538
Score:                  515.00         Matches:         515
Percent Similarity:     95.7%          Conservative:    0
Best Local Similarity:  95.7%          Mismatches:      23
Query Match:            21.0%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x AED10987 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60
```

```
Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet  80

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db       101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db       121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db       141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db       161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db       181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db       201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db       221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db       241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db       261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db       281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db       301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db       321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db       341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db       361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db       381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGlu***GlyGlyThrCysLeuAlaLeu 400

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db       401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
```

```
Db       421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db       441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db       481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy      7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db       501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy      7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db       521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 15
ADE09243
ID  ADE09243 standard; protein; 544 AA.
XX
AC  ADE09243;
XX
DT  29-JAN-2004  (first entry)
XX
DE  Novel protein-related contig polypeptide sequence #309.
XX
KW  novel gene; novel protein; tissue marker; molecular weight marker;
KW  chromosome marker; genetic disorder; contig.
XX
OS  Unidentified.
XX
PN  WO2003054152-A2.
XX
PD  03-JUL-2003.
XX
PF  10-DEC-2002; 2002WO-US039555.
XX
PR  10-DEC-2001; 2001US-0339739P.
PR  11-DEC-2001; 2001US-0339453P.
PR  14-MAR-2002; 2002US-0365091P.
PR  14-MAR-2002; 2002US-0365384P.
PR  12-APR-2002; 2002US-0372381P.
PR  12-APR-2002; 2002US-0372615P.
PR  22-APR-2002; 2002US-00128558.
PR  24-APR-2002; 2002US-0376045P.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Tang YT,  Asundi V,  Goodrich RW,  Ren F,  Zhang J,  Zhao QA,  Wang J;
PI  Ghosh M,  Xue AJ,  Wehrman T,  Weng G,  Zhou P,  Drmanac RT,  Wang Z;
PI  Ma Y,  Wang D,  Chen R,  Xu C,  Boyle BJ;
XX
DR  WPI; 2003-569235/53.
XX
PT  New polynucleotides, useful for expressing recombinant proteins for
PT  analysis, characterization or therapeutic use, or as markers for tissues
PT  in which the corresponding protein is preferentially expressed.
XX
PS  Disclosure; SEQ ID NO 2787; 1177pp; English.
XX
CC  The invention comprises the amino acid and coding sequences of novel
CC  proteins. The DNA and protein sequences of the invention are useful as:
CC  markers for tissues in which the corresponding protein is preferentially
CC  expressed; as molecular weight markers on gels; as chromosome markers or
CC  tags; to identify chromosomes or to map related gene positions; and to
CC  compare with endogenous DNA sequences in patients to identify potential

```
CC   genetic disorders. The present amino acid sequence was used in the
CC   exemplification of the invention.
XX
SQ   Sequence 544 AA;

Alignment Scores:
Pred. No.:              0.563          Length:         544
Score:                  500.50         Matches:        511
Percent Similarity:     93.9%          Conservative:   0
Best Local Similarity:  93.9%          Mismatches:     32
Query Match:            20.5%          Indels:         1
DB:                     1              Gaps:           1

US-10-717-580-11-COPY (1-7582) x ADE09243 (1-544)

Qy       5566 AATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCC 5625
Db          1 AsnHisLysSerProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuPro 20

Qy       5626 TSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMA 5685
Db         21 SerPheThrLeuThrAlaProProProCysArgCysMetThrSerSerSerProTyrGln 40

Qy       5686 GAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTT 5745
Db         41 GluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeu 60

Qy       5746 TSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCT 5805
Db         61 SerLysGlyThrProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSer 80

Qy       5806 GCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGT 5865
Db         81 AlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSer 100

Qy       5866 TGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCT 5925
Db        101 CysProGlyGlyLeuGlyValThrValCysArgThrTyrPheThrGlnThrGlyMetSer 120

Qy       5926 GATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAA 5985
Db        121 AspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLysGluAlaIleSerGln 140

Qy       5986 CTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACAT 6045
Db        141 LeuThrArgGlyHisSerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHis 160

Qy       6046 GAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTC 6105
Db        161 GluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeu 180

Qy       6106 CATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARG 6165
Db        181 HisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArg 200

Qy       6166 CCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 6225
Db        201 ProTyrValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThr 220

Qy       6226 ACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTC 6285
Db        221 ThrSerValLeuValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeu 240

Qy       6286 ACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTA 6345
Db        241 ThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpVal 260

Qy       6346 ACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCA 6405
Db        261 ThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSer 280

Qy       6406 GCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCC 6465
Db        281 AlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValPro 300

Qy       6466 CCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCT---AAGCCCCGC 6522
Db        301 ProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSer***SerProArg 320

Qy       6523 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 6582
Db        321 AsnLysArgValProIleLeuProPheValIleGlyAlaGlyValLeuGlyGlyLeuGly 340

Qy       6583 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 6642
Db        341 ThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyrHisLysLeuSerGlnGluLeu 360

Qy       6643 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
Db        361 AsnGlyAspMetGluGlnValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSer 380

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
Db        381 LeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGly 400

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 6822
Db        401 GlyThrCysLeuLeuLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleVal 420

Qy       6823 ACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAAC 6882
Db        421 ThrGluLysValLysGluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsn 440

Qy       6883 ACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCT 6942
Db        441 ThrGlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyPro 460

Qy       6943 CTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAAC 7002
Db        461 LeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsn 480

Qy       7003 TTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAG 7062
Db        481 PheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLys 500

Qy       7063 ACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
Db        501 ThrLysIleTyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAsp 520

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
Db        521 IleLysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSer 540

Qy       7183 GCAGGAAGCAGT 7194
Db        541 AlaGlySerSer 544

RESULT 16
ADF60170
ID   ADF60170 standard; protein; 544 AA.
XX
AC   ADF60170;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human contig polypeptide sequence SEQ ID NO:2537.
XX
KW   biological activity; genetic engineering; hybridisation probe; oligomer;
KW   primer; chromosome mapping; gene mapping; recombinant protein production;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO2003080795-A2.
XX
PD   02-OCT-2003.
```

XX
PF 09-AUG-2002; 2002WO-US025485.
XX
PR 09-AUG-2001; 2001US-0311261P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
DR WPI; 2003-876918/81.
DR N-PSDB; ADF59718.
XX
PT New polynucleotides, useful as hybridization probes, oligomers or
PT primers, for chromosome or gene mapping, for the recombinant production
PT of proteins, and for generating antisense DNA or RNA.
XX
PS Example 3; SEQ ID NO 2537; 571pp; English.
XX
CC The present invention describes isolated polynucleotide sequences (I),
CC which encode polypeptides (II) with biological activity. Also described:
CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
CC a host cell genetically engineered to comprise (I) which is operatively
CC associated with a regulatory sequence that modulates expression of (I) in
CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of
CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
CC polynucleotides comprising at least one of the polynucleotide sequences
CC (I). The polynucleotides (I) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA. The
CC present sequence represents a human contig polypeptide sequence, which is
CC used in an example from the present invention.
XX
SQ Sequence 544 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.563 | Length: | 544 |
| Score: | 500.50 | Matches: | 511 |
| Percent Similarity: | 93.9% | Conservative: | 0 |
| Best Local Similarity: | 93.9% | Mismatches: | 32 |
| Query Match: | 20.5% | Indels: | 1 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x ADF60170 (1-544)

```
Qy       5566 AATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCC 5625
             |||||||||||||||||  ||||||||||||||||||||      |||      |||
Db          1 AsnHisLysSerProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuPro 20

Qy       5626 TSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMA 5685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerPheThrLeuThrAlaProProProCysArgCysMetThrSerSerSerProTyrGln 40

Qy       5686 GAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTT 5745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeu 60

Qy       5746 TSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCT 5805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerLysGlyThrProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSer 80

Qy       5806 GCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGT 5865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSer 100

Qy       5866 TGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCT 5925
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        101 CysProGlyGlyLeuGlyValThrValCysArgThrTyrPheThrGlnThrGlyMetSer 120

Qy       5926 GATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAA 5985
              |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
```

```
Db        121 AspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLysGluAlaIleSerGln 140

Qy       5986 CTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACAT 6045
              ||||||   |||  |||  ||||||||||||||||||  ||||||||||||||||||||
Db        141 LeuThrArgGlyHisSerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHis 160

Qy       6046 GAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTC 6105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeu 180

Qy       6106 CATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARG 6165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArg 200

Qy       6166 CCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 6225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProTyrValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThr 220

Qy       6226 ACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTC 6285
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        221 ThrSerValLeuValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeu 240

Qy       6286 ACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTA 6345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpVal 260

Qy       6346 ACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCA 6405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSer 280

Qy       6406 GCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCC 6465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValPro 300

Qy       6466 CCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCT---AAGCCCCGC 6522
              ||||||  ||||||||||||||||||||||||||||||||||||||     ||||||
Db        301 ProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSer***SerProArg 320

Qy       6523 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 6582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AsnLysArgValProIleLeuProPheValIleGlyAlaGlyValLeuGlyGlyLeuGly 340

Qy       6583 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 6642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyrHisLysLeuSerGlnGluLeu 360

Qy       6643 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        361 AsnGlyAspMetGluGlnValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSer 380

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
              ||||||||||||||||||  ||||||||||||||||||  ||||||||||||||||||||
Db        381 LeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGly 400

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 6822
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyThrCysLeuLeuLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleVal 420

Qy       6823 ACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAAC 6882
              ||||||||||||  ||||||  ||||||||||||||  ||||||||||||||||||||||
Db        421 ThrGluLysValLysGluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsn 440

Qy       6883 ACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCT 6942
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        441 ThrGlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyPro 460

Qy       6943 CTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAAC 7002
              ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db        461 LeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsn 480

Qy       7003 TTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAG 7062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLys 500
```

```
Qy       7063 ACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        501 ThrLysIleTyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAsp 520

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 IleLysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSer 540

Qy       7183 GCAGGAAGCAGT 7194
              ||||||||||||
Db        541 AlaGlySerSer 544

RESULT 17
AAE25054
ID   AAE25054 standard; protein; 542 AA.
XX
AC   AAE25054;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Human EMBRY-2 protein.
XX
KW   Human; embryogenesis associated protein; AIDS; reproductive disorder;
KW   infertility; endometriosis; endometrial tumour; inflammatory disorder;
KW   autoimmune disorder; acquired immune deficiency syndrome; transgenic;
KW   ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
KW   EMBRY-2; allergy; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .20
FT                   /label= Signal_peptide
FT   Protein         21. .542
FT                   /note= "Mature EMBRY-2 protein"
FT   Domain          263. .484
FT                   /note= "ENV polyprotein (coat polyprotein) domain"
FT   Domain          455. .473
FT                   /note= "Transmembrane domain"
XX
PN   WO200248362-A2.
XX
PD   20-JUN-2002.
XX
PF   14-NOV-2001; 2001WO-US043956.
XX
PR   15-NOV-2000; 2000US-0249407P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Ramkumar J,  Arvizu C;
XX
DR   WPI; 2002-537629/57.
DR   N-PSDB; AAD41225.
XX
PT   New polypeptides of human embryogenesis associated proteins for screening
PT   modulators useful for treating or preventing disorders e.g.
PT   endometriosis, infertility, allergy, preeclampsia.
XX
PS   Claim 57; Page 94-95; 97pp; English.
XX
CC   The invention relates to human embryogenesis associated proteins (EMBRY)
CC   and nucleic acid molecules encoding such proteins. EMBRY sequences are
CC   useful for screening modulators useful for treating or preventing
CC   disorders associated with abnormal expression of EMBRY. The disorders
CC   treated include reproductive disorders such as infertility,
CC   endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
CC   disorder such as acquired immune deficiency syndrome (AIDS), allergies,
CC   contact dermatitis; disorders of the placenta such as preeclampsia,
CC   abruptio placentae etc. Sequences of the invention are also useful for
CC   analysing a proteome of a tissue or a cell type. EMBRY proteins are
CC   useful as immunogens for preparing antibodies. Polynucleotides of the
CC   invention are useful for creating knockin humanised animals or transgenic
CC   animals to model human diseases. They are also used in gene therapy. The
CC   present sequence is human EMBRY-2 protein
XX
SQ   Sequence 542 AA;

Alignment Scores:
Pred. No.:              1.09          Length:          542
Score:                  448.00        Matches:         460
Percent Similarity:     84.9%         Conservative:    0
Best Local Similarity:  84.9%         Mismatches:      78
Query Match:            18.3%         Indels:          4
DB:                     1             Gaps:            1

US-10-717-580-11-COPY (1-7582) x AAE25054 (1-542)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||   |||||||||       |||       |||   |||   ||||||
Db          1 MetAlaLeuProTyrCysIlePheLeuPheThrValLeuSerProProPheSerLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||   ||||||   |||   |||||||||||||||          ||||||||||||
Db         21 AlaProSerProCysHisCysArgThrSerSerSerProTyrGlnAlaPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||   ||||||       |||||||||||||||||||||||||||||   ||||||   |||
Db         41 MetArgArgProArgHisIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||   |||||||||   |||||||||||||||||||
Db         61 AlaPheThrAlaHisThrHisMetProHisAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||   |||   |||||||||||||||||   ||||||||||||||||||||||||||||
Db        101 GlyAlaThrIleCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerThrProSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrHisLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||   ||||||||||   |||   ||||||   |||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerVal 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerIleLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleIleHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAlaThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
```

```
                ||||||
```

```
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||  ||||||||||||  |||  ||||||  ||||||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||  ||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||      |||      ||||||  ||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||  ||||||||||||||  ||||||||||||||||||  |||
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                ||||||||||||||||  ||||||||  ||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              |||||||||  ||||||||||||||||||||  ||||||||||||||||  |||
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||  ||||||||||||      ||||||||||||  ||||||      |||||||||
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||  ||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpThrLeuProPheLeuGlyProLeuAlaAlaIleIlePhe 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||||      ||||||  ||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnPheLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAA------------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              ||||||||||||            ||||||||||||||  |||||||||  ||||||
Db        481 GluAlaValLysLeuGlnIleValLeuGlnMetGluProGlnMetGlnSerMetThrLys 500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              |||||||||  ||||||||||||||||            ||||||||||||||||||
.
```

```
Db        501 IleTyrArgGlyProLeuAspArgProAlaArgLeuCysSerAspValAsnAspIleGlu 520

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
                ||||||||||||||||||||  ||||||||||||||||      ||||||  |||
Db        521 ValThrProProGluGluIleSerThrAlaGlnProLeuLeuHisSerAsnSerValGly 540

Qy       7189 AGCAGT 7194
              ||||||
Db        541 SerSer 542

RESULT 19
AEB44408
ID   AEB44408 standard; protein; 542 AA.
XX
AC   AEB44408;
XX
DT   06-OCT-2005  (first entry)
XX
DE   MSRV/HERV-W ENV-pV14 protein sequence.
XX
KW   pharmaceutical; antibody; TLR-4; Toll-like receptor-4; neuroprotective;
KW   neuroleptic; multiple sclerosis; schizophrenia; inflammation;
KW   antiinflammatory; env protein; Env-pV14.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= Signal_peptide
FT   Protein         31. .542
FT                   /label= Mature_ENV_pV14
FT   Misc-difference 446
FT                   /label= OTHER
FT                   /note= "OTHER= Unidentified amino acid (unreadable in
FT                   specification)"
FT   Region          449. .542
FT                   /note= "This region contains the trans-membrane and intra
FT                   -cytoplasmic domain"
XX
PN   FR2865403-A1.
XX
PD   29-JUL-2005.
XX
PF   23-JAN-2004; 2004FR-00000675.
XX
PR   23-JAN-2004; 2004FR-00000675.
XX
PA   (INMR ) BIOMERIEUX SA.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Marche P,  Rolland A,  Jouvin ME,  Perron H;
XX
DR   WPI; 2005-514257/53.
XX
PT   Compound for treating pathologies connected with MSRV/HERV-W protein
PT   contains antibodies capable of bonding specifically with soluble
PT   fraction.
XX
PS   Disclosure; Fig 1; 51pp; French.
XX
CC   This invention relates to a novel composition comprising at least one
CC   antibody selected from anti-Env-SU MRSV/HERV-W or the group of anti-TLR4
CC   antibodies capable of bonding specifically with soluble fraction of the
CC   Env MRSV/HERV-W protein or the TLR4 receptor of the same protein, in a
CC   pharmaceutically acceptable vehicle or vector. The invention may be
CC   useful for the development of compounds with a neuroprotective or
CC   neuroleptic activity. The composition is useful for treating pathologies
CC   associated with MRSV/HERV-W such as multiple sclerosis and schizophrenia.
CC   The antibodies combat the inflammatory activity of the proteins causing
CC   the pathologies. The present sequence is that of the MRSV/HERV-W Env-p14
CC   protein from which the Env-Su protein, the soluble fragment of Env-pV14
CC   which was used during the development of the novel compositions of the
CC   invention, was obtained.
```

XX
SQ Sequence 542 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 1.22 | Length: | 542 |
| Score: | 439.00 | Matches: | 451 |
| Percent Similarity: | 83.2% | Conservative: | 0 |
| Best Local Similarity: | 83.2% | Mismatches: | 87 |
| Query Match: | 17.9% | Indels: | 4 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x AEB44408 (1-542)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMet***TrpThrLeuProArgLeuGlyProLeuAlaAlaIleIlePhe 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnPheLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAA----------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
Db        481 GluAlaValLysLeuGlnIleValLeuGlnMetGluProGlnMetGlnSerMetThrLys 500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
Db        501 IleTyrArgGlyProLeuAspArgProAlaArgLeuCysSerAspValAsnAspIleGlu 520

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
Db        521 ValThrProProGluGluIleSerThrAlaGlnProLeuLeuHisSerAsnSerValGly 540

Qy       7189 AGCAGT 7194
Db        541 SerSer 542
```

RESULT 20
AAB08195
ID AAB08195 standard; protein; 541 AA.
XX
AC AAB08195;
XX
DT 06-AUG-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
DE Amino acid sequence of the MSRV-1 3' env and LTR regions.
XX
KW MSRV-1; pol region; long terminal repeat; LTR; RU5 region; retrovirus.
XX
OS Multiple sclerosis associated retrovirus.
XX

| FH | Key | Location/Qualifiers |
|---|---|---|
| FT | Peptide | 1. .27 |
| FT | | /note= "signal peptide" |
| FT | Modified-site | 58. .60 |
| FT | | /note= "N-linked glycosylation site" |
| FT | Modified-site | 92. .94 |

```
FT                  /note= "N-linked glycosylation site"
FT   Modified-site  213. .215
FT                  /note= "N-linked glycosylation site"
FT   Modified-site  233. .235
FT                  /note= "N-linked glycosylation site"
FT   Modified-site  280. .282
FT                  /note= "N-linked glycosylation site"
FT   Cleavage-site  289. .290
FT                  /note= "putative cleavage site"
FT   Cleavage-site  315. .316
FT                  /note= "putative cleavage site"
FT   Modified-site  408. .410
FT                  /note= "N-linked glycosylation site"
XX
PN   WO200047745-A1.
XX
PD   17-AUG-2000.
XX
PF   15-FEB-2000; 2000WO-IB000159.
XX
PR   15-FEB-1999;   99EP-00420041.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Perron H,  Komurian-Pradel F;
XX
DR   WPI; 2000-506097/45.
DR   N-PSDB; AAA63826.
XX
PT   Nucleotide fragment of LTR-RU5 region from Multiple Sclerosis retrovirus
PT   (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
PT   sample.
XX
PS   Disclosure; Fig 2; 23pp; English.
XX
CC   The present sequence is encoded by a nucleotide sequence corressponding
CC   to the 3' env region and long terminal repeat sequences from clone CL6 of
CC   Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
CC   long terminal repeat (LTR)-RU5 region which encodes the expression of a
CC   MSRV-1 protein. This is unusual for LTRs, in particular in the RU5
CC   region. Probes and antibodies to the MSRV-1 retrovirus protein and
CC   encoding polynucleotide sequences are used to detect the presence of MSRV
CC   -1 retrovirus in a biological sample. (Updated on 06-AUG-2003 to correct
CC   OS field.)
XX
SQ   Sequence 541 AA;

Alignment Scores:
Pred. No.:              1.31          Length:        541
Score:                  433.50        Matches:       456
Percent Similarity:     84.1%         Conservative:  0
Best Local Similarity:  84.1%         Mismatches:    81
Query Match:            17.7%         Indels:        5
DB:                     1             Gaps:          2

US-10-717-580-11-COPY (1-7582) x AAB08195 (1-541)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||||   ||||||      |||      |||   |||   ||||||
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||   |||   ||||||||||||||      |||||||||   |||
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu---Arg 39

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
                  |||||||||||||||||||||||||||||||||||||||   ||||||
Db         40 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 59

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db         60 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 79

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 99

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||   |||||||||||||||||||||||   |||   ||||||||||||||||||
Db        100 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 119

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||   ||||||||||||||   |||||||||   ||||||||||||||       |||
Db        120 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 139

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                  |||   ||||||||||||||||||   |||||||||||||||||||||||||||||
Db        140 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 159

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        160 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 179

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||   ||||||||||||   |||   ||||||   ||||||
Db        180 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 199

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 219

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        220 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 239

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||      ||||||   ||||||||||||||||||||||||||||||||
Db        240 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 259

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 279

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 299

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||   |||   ||||||   ||||||||||||||||||||
Db        300 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 319

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||   |||
Db        320 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 339

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        340 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 359

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 379

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              |||||||||||||||||||   |||||||||   ||||||||||||||||||||||||||
Db        380 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 399

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              |||||||||   ||||||||||||||||||||   ||||||||||||||||||   |||
Db        400 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 419

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||      ||||||||||||   ||||||      |||||||||
Db        420 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 439

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
```

```
                ||||||   ||||||||||||   ||||||||||||||||||||||||||||   ||||||
Db        440 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 459

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||||   |||||||||   ||||||||||||||||||
Db        460 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 479

Qy       7021 GAAGCTGTRAAA------------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
             ||||||||||||            ||||||||||||||   |||   |||   ||||||
Db        480 GluAlaValLysLeuGlnMetValLeuGlnMetGluProGlnMetGluSerMetThrLys 499

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
             |||   |||   ||||||||||||||||   ||||||   ||||||||||||||||||
Db        500 IleHisArgGlyProLeuAspArgProAlaSerProCysSerAspValAsnAspIleGlu 519

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
             ||||||||||||||||||||||   ||||||||||||||   ||||||||||||||||
Db        520 GlyThrProProGluGluIleSerThrAlaGlnProLeuLeuCysProAsnSerAlaGly 539

Qy       7189 AGCAGT 7194
             ||||||
Db        540 SerSer 541

RESULT 21
AAW71068
ID   AAW71068 standard; protein; 493 AA.
XX
AC   AAW71068;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus protein 3.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 39
FT                   /note= "Encoded by tga"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   N-PSDB; AAV43217.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 185-187; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) protein used in the method of the invention. The
CC   invention provides complete or partial genomic sequences of the MSRV-1
CC   pol gene, gag gene and env gene, and polypeptides encoded by these genes.
CC   The invention also provides antibodies raised against the polypeptides.
CC   The genomic sequences, polypeptides and antibodies are also claimed
CC   useful for diagnosing infection by MS and rheumatoid arthritis-associated
CC   viruses, and also for prevention and treatment of infection with these
CC   viruses
XX
SQ   Sequence 493 AA;

Alignment Scores:
Pred. No.:              1.63          Length:         493
Score:                  416.00        Matches:        416
Percent Similarity:     84.7%         Conservative:   0
Best Local Similarity:  84.7%         Mismatches:     75
Query Match:            17.0%         Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x AAW71068 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   ||||||||||||   ||||||      |||       |||   |||   ||||||
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||   |||   ||||||||||||||       |||||||||   |||
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
                  ||||||||||||||||||||||||||||||||||   ||||||
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||   ||||||||||||||||||||||   |||   ||||||||||||||||||
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||   ||||||||||||||   |||||||||   ||||||||||||||   |||
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                 |||   ||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||   ||||||||||||   |||   ||||||   ||||||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||   ||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
```

```
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 22
AAW99553
ID   AAW99553 standard; protein; 493 AA.
XX
AC   AAW99553;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Protein encoded by clone C15 from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis.
XX
OS   Viruses.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 39
FT                   /label= unknown
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   N-PSDB; AAX29703.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 2; Page 38-39; 83pp; French.
XX
CC   This sequence represents the protein encoded by clone C15 from a novel
CC   multiple sclerosis related virus type 1 (MSRV1). The sequence can be used
CC   in diagnostic, prophylactic or therapeutic compositions to inhibit
CC   expression of a multiple sclerosis related virus and/or virus associated
CC   with rheumatoid polyarthritis. (Updated on 27-AUG-2003 to correct OS
CC   field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 493 AA;

Alignment Scores:
Pred. No.:              1.63          Length:          493
Score:                  416.00        Matches:         416
Percent Similarity:     84.7%         Conservative:    0
Best Local Similarity:  84.7%         Mismatches:      75
Query Match:            17.0%         Indels:          0
DB:                     1             Gaps:            0

US-10-717-580-11-COPY (1-7582) x AAW99553 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
```

```
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 23
ADG14847
ID   ADG14847 standard; protein; 493 AA.
XX
AC   ADG14847;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated protein #7.
XX
KW   pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 106; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 493 AA;

Alignment Scores:
Pred. No.:              1.63          Length:          493
Score:                  416.00        Matches:         416
Percent Similarity:     84.7%         Conservative:    0
Best Local Similarity:  84.7%         Mismatches:      75
Query Match:            17.0%         Indels:          0
DB:                     1             Gaps:            0

US-10-717-580-11-COPY (1-7582) x ADG14847 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200
```

```
Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||   ||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||      |||      ||||||   ||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||   ||||||||||||||   ||||||||||||||||||   |||
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                ||||||||||||||||   |||||||||   ||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              |||||||||   ||||||||||||||||||||   ||||||||||||||||||   |||
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||      ||||||||||||   ||||||      |||||||||
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||   ||||||||||||||||||||||||   ||||||
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||   |||||||||   ||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
              ||||||||||||||||||||         |||
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 24
ADB84401
ID   ADB84401 standard; protein; 493 AA.
XX
AC   ADB84401;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated protein/peptide sequence #10.
XX
KW   multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Claim 31; Page 78-80; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is a claimed MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 493 AA;

Alignment Scores:
Pred. No.:              1.74          Length:         493
Score:                  411.00        Matches:        411
Percent Similarity:     83.7%         Conservative:   0
Best Local Similarity:  83.7%         Mismatches:     80
Query Match:            16.8%         Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x ADB84401 (1-493)

Qy     c  5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
               |||   ||||||||||||   ||||||      |||      |||   |||   ||||||
Db           1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20
```

```
Qy      5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db        21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy      5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db        41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db       101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db       121 GlnGlyGlnAlaArgGluLys***ValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db       141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db       161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db       181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db       201 ProValProGluGlnTrpAsn***PheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db       221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db       241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db       261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db       281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db       301 ThrGluGln***LeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db       321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThr***IleGlySerIle 340

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db       341 ThrThrSerThrGlnPheTyrTyrLysLeuSer***GluIleAsnGlyAspMetGluGln 360

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db       361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db       381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db       401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThr***LysValLysGlu 420

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db       421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db       441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db       481 GluAlaValLysLeuGlnMetValLeuGlnMet 491
```

RESULT 25
AEB44407
ID AEB44407 standard; protein; 287 AA.
XX
AC AEB44407;
XX
DT 06-OCT-2005 (first entry)
XX
DE MSRV/HERV-W ENV-pSU protein sequence SeqID1.
XX
KW pharmaceutical; antibody; TLR-4; Toll-like receptor-4; neuroprotective;
KW neuroleptic; multiple sclerosis; schizophrenia; inflammation;
KW antiinflammatory; env protein; Env-SU.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN FR2865403-A1.
XX
PD 29-JUL-2005.
XX
PF 23-JAN-2004; 2004FR-00000675.
XX
PR 23-JAN-2004; 2004FR-00000675.
XX
PA (INMR ) BIOMERIEUX SA.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Marche P, Rolland A, Jouvin ME, Perron H;
XX
DR WPI; 2005-514257/53.
XX
PT Compound for treating pathologies connected with MSRV/HERV-W protein
PT contains antibodies capable of bonding specifically with soluble
PT fraction.
XX
PS Disclosure; SEQ ID NO 1; 51pp; French.
XX
CC This invention relates to a novel composition comprising at least one
CC antibody selected from anti-Env-SU MRSV/HERV-W or the group of anti-TLR4
CC antibodies capable of bonding specifically with soluble fraction of the
CC Env MRSV/HERV-W protein or the TLR4 receptor of the same protein, in a
CC pharmaceutically acceptable vehicle or vector. The invention may be
CC useful for the development of compounds with a neuroprotective or
CC neuroleptic activity. The composition is useful for treating pathologies
CC associated with MRSV/HERV-W such as multiple sclerosis and schizophrenia.
CC The antibodies combat the inflammatory activity of the proteins causing
CC the pathologies. The present sequence is that of the MRSV/HERV-W Env-SU
CC protein (a portion of the Env pV14 protein) which was used during the
CC development of the novel compositions of the invention.
XX

```
SQ   Sequence 287 AA;

Alignment Scores:
Pred. No.:              13          Length:         287
Score:                  249.00      Matches:        249
Percent Similarity:     86.8%       Conservative:   0
Best Local Similarity:  86.8%       Mismatches:     38
Query Match:            10.2%       Indels:         0
DB:                     1           Gaps:           0

US-10-717-580-11-COPY (1-7582) x AEB44407 (1-287)

Qy       5668 AGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCC 5727
                ||||||||||||      ||||||||||||||||          ||||||||||||||||||
Db          1 SerSerSerProTyrGlnGluPheLeuTrpArgThrArgLeuProGlyAsnIleAspAla 20

Qy       5728 CCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCC 5787
                |||||||||||||||||    ||||||      ||||||||||||||||||||||||||||
Db         21 ProSerTyrArgSerLeuSerLysGlyAsnSerThrPheThrAlaHisThrHisMetPro 40

Qy       5788 CGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGA 5847
                ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgAsnCysTyrAsnSerAlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGly 60

Qy       5848 AAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTC 5907
                ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         61 LysMetIleAsnProSerCysProGlyGlyLeuGlyAlaThrValCysTrpThrTyrPhe 80

Qy       5908 ACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTA 5967
                |||   |||   ||||||||||||||||||   |||   ||||||||||||||||   |||
Db         81 ThrHisThrSerMetSerAspGlyGlyGlyIleGlnGlyGlnAlaArgGluLysGlnVal 100

Qy       5968 AAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTA 6027
                ||||||   ||||||||||||||||      |||   |||   ||||||||||||||||||
Db        101 LysGluAlaIleSerGlnLeuThrArgGlyHisSerThrProSerProTyrLysGlyLeu 120

Qy       6028 GATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAAT 6087
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ValLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsn 140

Qy       6088 ACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGC 6147
                ||||||||||||   ||||||||||||||||||||||||||||||||||||   |||
Db        141 ThrThrLeuThrArgLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpMetCys 160

Qy       6148 CTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTC 6207
                |||||||||   |||   ||||||   ||||||||||||||||||||||||||||||||
Db        161 LeuProLeuHisPheArgProTyrIleSerIleProValProGluGlnTrpAsnAsnPhe 180

Qy       6208 AGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATA 6267
                ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        181 SerThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsnLeuGluIle 200

Qy       6268 ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCC 6327
                ||||||||||||||||||||||||||||||||||||||      ||||||   |||
Db        201 ThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAspThrThrSerSer 220

Qy       6328 CAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTT 6387
                ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db        221 GlnCysIleArgTrpValThrProProThrArgIleValCysLeuProSerGlyIlePhe 240

Qy       6388 TTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTC 6447
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGluSerMetCysPhe 260

Qy       6448 CTCTCATTCTTAGTGCCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTC 6507
                |||||||||||||||||||||||||    ||||||||||||||||||||||||      |||
Db        261 LeuSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyrAsnHisVal 280

Qy       6508 ATATCTAAGCCCCGCAACAAA 6528
                ||||||   ||||||
Db        281 ValProLysProHisAsnLys 287

RESULT 26
```

```
AAM18008
ID   AAM18008 standard; protein; 263 AA.
XX
AC   AAM18008;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Peptide #4442 encoded by probe for measuring cervical gene expression.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 27; SEQ ID NO 22834; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC   by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC   can be used to produce a single exon microarray, which can be used for
CC   measuring human gene expression in a sample derived from human cervical
CC   epithelial cells. By measuring gene expression, the probes are therefore
CC   useful in grading and/or staging of diseases of the cervix, notably
CC   cervical cancer. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1        Length:         263
Score:                  221.50      Matches:        232
Percent Similarity:     87.9%       Conservative:   0
Best Local Similarity:  87.9%       Mismatches:     31
Query Match:            9.1%        Indels:         1
DB:                     1           Gaps:           1

US-10-717-580-11-COPY (1-7582) x AAM18008 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
                ||||||||   ||||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60
```

```
Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
             ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
             |||   ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
             ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
             |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
             |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
             ||||||||||||||||        ||||||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
             ||||||||||||||||||||||||||||||||||||||||   |||||||||   |||||||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
             ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
             ||||||   |||
Db        260 LeuAspLeuLeu 263
```

RESULT 27
ABB37043
ID  ABB37043 standard; peptide; 263 AA.
XX
AC  ABB37043;
XX
DT  04-FEB-2002  (first entry)
XX
DE  Peptide #4549 encoded by human foetal liver single exon probe.
XX
KW  Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS  Homo sapiens.
XX
PN  WO200157277-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US000669.
XX
PR  04-FEB-2000; 2000US-0180312P.
PR  26-MAY-2000; 2000US-0207456P.
PR  30-JUN-2000; 2000US-00608408.
PR  03-AUG-2000; 2000US-00632366.
PR  21-SEP-2000; 2000US-0234687P.
PR  27-SEP-2000; 2000US-0236359P.
PR  04-OCT-2000; 2000GB-00024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR  WPI; 2001-483447/52.
XX
PT  Human genome-derived single exon nucleic acid probes useful for analyzing
PT  gene expression in human fetal liver.
XX
PS  Claim 27; SEQ ID NO 29678; 639pp + Sequence Listing; English.
XX
CC  The invention relates to a single exon nucleic acid probe for measuring
CC  human gene expression in a sample derived from human foetal liver. The
CC  single exon nucleic acid probes may be used for predicting, measuring and
CC  displaying gene expression in samples derived from human fetal liver. The
CC  present sequence is a peptide encoded by a single exon nucleic acid probe
CC  of the invention. Note: The sequence data for this patent did not form
CC  part of the printed specification, but was obtained in electronic format
CC  directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 263 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 18.1 | Length: | 263 |
| Score: | 221.50 | Matches: | 232 |
| Percent Similarity: | 87.9% | Conservative: | 0 |
| Best Local Similarity: | 87.9% | Mismatches: | 31 |
| Query Match: | 9.1% | Indels: | 1 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x ABB37043 (1-263)

```
Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
             |||||||||  ||||||||||||||||||||||||  ||||||  |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
             ||||||||||||||||||||||||||||||||||||||||||  |||||||||  ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
             ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
             |||   ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
             ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
             |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
             |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
```

```
                ||||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
                ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
                ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 28
AAM30519
ID   AAM30519 standard; protein; 263 AA.
XX
AC   AAM30519;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Peptide #4556 encoded by probe for measuring placental gene expression.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 27; SEQ ID NO 30788; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP:
CC   see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for producing a microarray for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from human placenta. The probes are useful for antenatal diagnosis of
CC   human genetic disorders
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:         263
Score:                  221.50        Matches:        232
Percent Similarity:     87.9%         Conservative:   0
Best Local Similarity:  87.9%         Mismatches:     31
Query Match:            9.1%          Indels:         1
DB:                     1             Gaps:           1

US-10-717-580-11-COPY (1-7582) x AAM30519 (1-263)
```

```
Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
                ||||||||   ||||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
                ||||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db          21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
                ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db          61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db          81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
                |||   ||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
                ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
                |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db         160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
                |||      ||||||   ||||||||||||||||||||||||||||||||||||||||
Db         180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
                ||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db         200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db         220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
                ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
                ||||||   |||
Db         260 LeuAspLeuLeu 263

RESULT 29
ABB31812
ID   ABB31812 standard; peptide; 263 AA.
XX
AC   ABB31812;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Peptide #4463 encoded by breast cell single exon nucleic acid probe.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
```

```
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 27; SEQ ID NO 14780; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a peptide
CC   encoded by a single exon nucleic acid probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:         263
Score:                  221.50        Matches:        232
Percent Similarity:     87.9%         Conservative:   0
Best Local Similarity:  87.9%         Mismatches:     31
Query Match:            9.1%          Indels:         1
DB:                     1             Gaps:           1

US-10-717-580-11-COPY (1-7582) x ABB31812 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
              |||||||||   |||||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
              |||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
              ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              |||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              |||   |||||||||||||||||||||||||||||||||   |||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              ||||||||||||||||||        ||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              |||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 30
ABB22358
ID   ABB22358 standard; protein; 263 AA.
XX
AC   ABB22358;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Protein #4357 encoded by probe for measuring heart cell gene expression.
XX
KW   Human; gene expression; heart; microarray; vascular system;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
```

```
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 15; SEQ ID NO 24128; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart (see
CC   ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC   probe. The probes may be used for predicting, measuring and displaying
CC   gene expression in samples derived from the human heart via microarrays.
CC   By measuring gene expression, the probes are useful for predicting,
CC   diagnosing, grading, staging, monitoring and prognosing diseases of the
CC   human heart and vascular system e.g. cardiovascular disease,
CC   hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:        263
Score:                  221.50        Matches:       232
Percent Similarity:     87.9%         Conservative:  0
Best Local Similarity:  87.9%         Mismatches:    31
Query Match:            9.1%          Indels:        1
DB:                     1             Gaps:          1

US-10-717-580-11-COPY (1-7582) x ABB22358 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
              |||||||||   |||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
              |||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
              ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              |||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              |||   |||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              |||   ||||||||||||||||||   ||||||   ||||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              ||||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              |||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 31
AAM70185
ID   AAM70185 standard; protein; 263 AA.
XX
AC   AAM70185;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed probe encoded protein SEQ ID NO: 30491.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 30491; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is a
CC   protein encoded by one of the probes of the invention
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:        263
Score:                  221.50        Matches:       232
Percent Similarity:     87.9%         Conservative:  0
Best Local Similarity:  87.9%         Mismatches:    31
Query Match:            9.1%          Indels:        1
DB:                     1             Gaps:          1

US-10-717-580-11-COPY (1-7582) x AAM70185 (1-263)
```

```
Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
Db        260 LeuAspLeuLeu 263
```

RESULT 32
AAM57770
ID AAM57770 standard; protein; 263 AA.
XX
AC AAM57770;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29875.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.

XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 29875; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 263 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 18.1 | Length: | 263 |
| Score: | 221.50 | Matches: | 232 |
| Percent Similarity: | 87.9% | Conservative: | 0 |
| Best Local Similarity: | 87.9% | Mismatches: | 31 |
| Query Match: | 9.1% | Indels: | 1 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x AAM57770 (1-263)


```
Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
```

```
                ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
                |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
                |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
                ||||||||||||||||         ||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
                ||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
                ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
                ||||||   |||
Db        260 LeuAspLeuLeu 263
```

RESULT 33
ABG51885
ID ABG51885 standard; peptide; 263 AA.
XX
AC ABG51885;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID No 30533.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 30533; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 263 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 18.1 | Length: | 263 |
| Score: | 221.50 | Matches: | 232 |
| Percent Similarity: | 87.9% | Conservative: | 0 |
| Best Local Similarity: | 87.9% | Mismatches: | 31 |
| Query Match: | 9.1% | Indels: | 1 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x ABG51885 (1-263)

```
Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
                |||||||||   ||||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
                ||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
                ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
                |||   ||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
                ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
                |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
                |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
                ||||||||||||||||         ||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
                ||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
                ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259
```

```
Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 34
AAM05648
ID   AAM05648 standard; protein; 263 AA.
XX
AC   AAM05648;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Peptide #4330 encoded by probe for measuring breast gene expression.
XX
KW   Probe; human; breast disease; breast cancer; development disorder;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 27; SEQ ID NO 14388; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes
CC   (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for measuring human gene expression in
CC   a human breast sample, where the probe hybridises at high stringency to a
CC   nucleic acid expressed in the human breast. The probes are useful for
CC   predicting, diagnosing, grading, staging, monitoring and prognosing
CC   diseases of the human breast, particularly those diseases with polygenic
CC   aetiology. The diseases include: breast cancer, disorders of development,
CC   inflammatory diseases of the breast, fibrocystic changes, proliferative
CC   breast disease and non-carcinoma tumours. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:         263
Score:                  221.50        Matches:        232
Percent Similarity:     87.9%         Conservative:   0
Best Local Similarity:  87.9%         Mismatches:     31
Query Match:            9.1%          Indels:         1
DB:                     1             Gaps:           1

US-10-717-580-11-COPY (1-7582) x AAM05648 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
              |||||||||   ||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
              ||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db          21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
              ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db          61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db          81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              |||   ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db         120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db         140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||||
Db         160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              ||||||||||||||||||        ||||||||||||||||||||||||||||||||||||
Db         200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              ||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db         220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db         260 LeuAspLeuLeu 263

RESULT 35
ABG39819
ID   ABG39819 standard; peptide; 263 AA.
XX
AC   ABG39819;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human peptide encoded by genome-derived single exon probe SEQ ID 29484.
XX
KW   Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 27; SEQ ID NO 29484; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a peptide/protein encoded by a single exon probe of
CC   the invention. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 263 AA;

Alignment Scores:
Pred. No.:              18.1          Length:        263
Score:                  221.50        Matches:       232
Percent Similarity:     87.9%         Conservative:  0
Best Local Similarity:  87.9%         Mismatches:    31

Query Match:            9.1%          Indels:        1
DB:                     1             Gaps:          1

US-10-717-580-11-COPY (1-7582) x ABG39819 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
              ||||||||   ||||||||||||||||||||||   ||||||   |||      |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
              ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              |||   ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              ||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 36
AAB75137
ID   AAB75137 standard; protein; 120 AA.
XX
AC   AAB75137;
XX
DT   08-AUG-2001  (first entry)
XX
DE   HERV-W truncated envelope protein GT.
XX
KW   Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW   envelope protein; multiple sclerosis-related superantigen; vaccine;
```

```
KW   surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW   antisense-therapy; autoimmune disorder.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200131021-A1.
XX
PD   03-MAY-2001.
XX
PF   30-OCT-2000; 2000WO-EP010659.
XX
PR   28-OCT-1999;   99EP-00402690.
XX
PA   (UYGE-) UNIV GENEVE.
XX
PI   Conrad B,  Mach B;
XX
DR   WPI; 2001-316336/33.
XX
PT   New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT   activity useful for diagnosing and treating multiple sclerosis.
XX
PS   Claim 1; Fig 7; 94pp; English.
XX
CC   On the basis of the PBS t-RNA motif used for the classification of human
CC   endogenous retrovirus (HERVs) the full length endogenous provirus which
CC   was been located on the long arm of human chromosome 7 (7q21-22) has been
CC   designated HERV-W. The present invention describes proteins or peptides
CC   (I) having superantigen (SAg) activity comprising the ENV protein (ENV)
CC   of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC   have neuroprotective activity, and can be used in: vaccines; antisense-
CC   therapy; and HERV-W SAg activity-inhibitors. (I) and encoding DNA/RNA are
CC   useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC   disorders. (I) are also useful for identifying substances (and optionally
CC   recovering) capable of binding to a retroviral superantigen associated
CC   with MS, substances capable of blocking SAg activity and substances
CC   capable of blocking transcription or translation of HERV-W retroviral
CC   superantigen. A protein or peptide derived from (I), modified to be
CC   devoid of SAg activity and being capable of generating an immune response
CC   against HERV-W retroviral SAg is useful in therapy. Nucleic acid
CC   molecules encoding (I) are useful as vaccines against MS. Substances
CC   capable of blocking SAg activity, capable of binding to a retroviral
CC   superantigen associated with MS, or capable of blocking transcription or
CC   translation of HERV-W retroviral superantigen for use in treating or
CC   preventing MS, obtained using (I) are useful for the treatment and
CC   prevention of MS. (I) and nucleic acids encoding them are useful for
CC   diagnosing autoimmune disease. The present sequence represents the
CC   specifically claimed truncated envelope protein of HERV-W designated GT
XX
SQ   Sequence 120 AA;

Alignment Scores:
Pred. No.:              68.1          Length:         120
Score:                  111.00        Matches:        111
Percent Similarity:     92.5%         Conservative:   0
Best Local Similarity:  92.5%         Mismatches:     9
Query Match:            4.5%          Indels:         0
DB:                     1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x AAB75137 (1-120)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   ||||||||||||||||||||      |||     |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

RESULT 37
ADK35467
ID   ADK35467 standard; protein; 122 AA.
XX
AC   ADK35467;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Novel human polypeptide SeqID7549.
XX
KW   antiarthritic; antiparkinsonian; neuroprotective; nootropic;
KW   immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
KW   antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
KW   arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
KW   psoriasis; inflammatory bowel disease; infection; bacteria; virus;
KW   fungus; parasite; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .122
FT                   /label= OTHER
FT                   /note= "OTHER= All Xaa's in this sequence are unknown
FT                   amino acids or the site of a stop codon within the DNA
FT                   sequence"
XX
PN   WO200216439-A2.
XX
PD   28-FEB-2002.
XX
PF   05-MAR-2001; 2001WO-US004941.
XX
PR   07-MAR-2000; 2000US-00519705.
PR   19-MAY-2000; 2000US-00574454.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2002-280918/32.
XX
PT   Isolated polynucleotide encoding bone marrow derived polypeptides useful
PT   for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
PT   disease, and inflammatory bowel disease.
XX
PS   Claim 20; SEQ ID NO 7549; 504pp; English.
XX
CC   This invention relates to a novel isolated polynucleotide comprising a
CC   nucleotide sequence selected from one of 1680 sequences, a mature protein
CC   coding portion of them, an active domain of them and their complementary
CC   sequences. The invention may be useful for the production of compounds
CC   with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
CC   immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,
CC   antibacterial, antiviral, antifungal or antiparasitic activity. In
CC   addition, the disclosed sequences may be useful for gene therapy. The
CC   polypeptides or their antibodies are useful for treating many diseases
CC   such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
CC   psoriasis, inflammatory bowel disease and infections caused by bacteria,
CC   viruses, fungi or parasites. The present sequence is that of a human
CC   polypeptide of the invention.
XX
```

SQ Sequence 122 AA;

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 74.9 | Length: | 122 |
| Score: | 102.00 | Matches: | 102 |
| Percent Similarity: | 83.6% | Conservative: | 0 |
| Best Local Similarity: | 83.6% | Mismatches: | 20 |
| Query Match: | 4.2% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x ADK35467 (1-122)

```
Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db          1 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 20

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 40

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                 |||||||||||||||||   |||||||||   ||||||||||||||||||||||||||
Db         41 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 60

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||   ||||||||||||||||||   ||||||||||||||||||   |||
Db         61 GlyGluGluCysTrpTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 80

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   |||||||||         ||||||||||||   ||||||      |||||||||
Db         81 IleArgAspArgIleProCysArgAlaGluGluLeuGlnAsnThrGluHisTrpGlyLeu 100

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||   ||||||||||||   ||||||||||||   ||||||
Db        101 LeuSerGlnTrpMetProTrpValLeuProPheLeu***ProLeuAlaAlaLeuIleLeu 120

Qy       6961 CTACTC 6966
              ||||||
Db        121 LeuLeu 122
```

Search completed: August 4, 2006, 17:15:00
Job time : 64 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 4, 2006, 17:11:28 ; Search time 0.1 Seconds
(without alignments)
12.328 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 2447
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: UNITARY_AA
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1 seqs, 542 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Published_Applications_AA_New

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=new.seq -DB=aapubnewdb -SUFFIX=pto
-OUT=aapubnew.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1
-END=-1 -MATRIX=unitary_AA -TRANS=human40.cdi -LIST=1 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : aapubnewdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 441 | 18.0 | 542 | 1 | US-11-261-673-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-11-261-673-2
; Sequence 2, Application US/11261673
; Publication No. US20060088820A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, Herve
; APPLICANT: LAFONT, Monique
; TITLE OF INVENTION: Detecting superantigen activity in a biological sample
; FILE REFERENCE: 110631
; CURRENT APPLICATION NUMBER: US/11/261,673
; CURRENT FILING DATE: 2005-10-31
; PRIOR APPLICATION NUMBER: US/09/936,835
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/FR00/00691
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: FR 99/03622
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: FR 99/13755
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-261-673-2

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0 | Length: | 542 |
| Score: | 441.00 | Matches: | 453 |
| Percent Similarity: | 83.6% | Conservative: | 0 |
| Best Local Similarity: | 83.6% | Mismatches: | 85 |
| Query Match: | 18.0% | Indels: | 4 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x US-11-261-673-2 (1-542)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260
```

```
Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||      |||      ||||||   ||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||   ||||||||||||||   ||||||||||||||||||   |||
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                 ||||||||||||||||   |||||||||   ||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||   ||||||||||||||||||||   ||||||||||||||||   |||
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||      ||||||||||||   ||||||      |||||||||
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpThrLeuProPheLeuGlyProLeuAlaAlaIleIlePhe 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||      ||||||   ||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnPheLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAA----------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              ||||||||||||          ||||||||||||||   |||||||||   ||||||
Db        481 GluAlaValLysLeuGlnIleValLeuGlnMetGluProGlnMetGlnSerMetThrLys 500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              |||||||||   ||||||||||||||||            ||||||||||||||||||
Db        501 IleTyrArgGlyProLeuAspArgProAlaArgLeuCysSerAspValAsnAspIleGlu 520

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
                 ||||||||||||||||||||   ||||||||||||||      ||||||   |||
Db        521 ValThrProProGluGluIleSerThrAlaGlnProLeuLeuHisSerAsnSerValGly 540

Qy       7189 AGCAGT 7194
              ||||||
Db        541 SerSer 542
```

Search completed: August 4, 2006, 17:11:32
Job time : 2.5 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 4, 2006, 17:09:08 ; Search time 2.4 Seconds
(without alignments)
7.530 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 2447
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: UNITARY_AA
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16 seqs, 7945 residues

Total number of hits satisfying chosen parameters: 32

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 16 summaries

Published_Applications_AA_Main

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=new.seq -DB=aapubmaindb -SUFFIX=pto
-OUT=aapubmain.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1
-END=-1 -MATRIX=unitary_AA -TRANS=human40.cdi -LIST=16 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=16 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : aapubmaindb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 641 | 26.2 | 685 | 1 | US-11-028-539-35 | Sequence 35, Appl |
| 2 | 518 | 21.2 | 540 | 1 | US-10-821-234-1395 | Sequence 1395, Ap |
| 3 | 518 | 21.2 | 540 | 1 | US-11-028-539-26 | Sequence 26, Appl |
| 4 | 517 | 21.1 | 538 | 1 | US-10-133-036-2 | Sequence 2, Appli |
| 5 | 517 | 21.1 | 538 | 1 | US-10-133-036-20 | Sequence 20, Appl |
| 6 | 517 | 21.1 | 538 | 1 | US-10-133-036-21 | Sequence 21, Appl |
| 7 | 517 | 21.1 | 538 | 1 | US-10-717-580-33 | Sequence 33, Appl |
| 8 | 516 | 21.1 | 538 | 1 | US-09-902-535-2 | Sequence 2, Appli |
| 9 | 516 | 21.1 | 538 | 1 | US-10-114-893-135 | Sequence 135, App |
| 10 | 516 | 21.1 | 538 | 1 | US-10-016-249-4 | Sequence 4, Appli |
| 11 | 448 | 18.3 | 542 | 1 | US-10-416-642-2 | Sequence 2, Appli |
| 12 | 418 | 17.1 | 503 | 1 | US-10-637-565-20 | Sequence 20, Appl |
| 13 | 416 | 17.0 | 493 | 1 | US-08-979-847-106 | Sequence 106, App |
| 14 | 416 | 17.0 | 493 | 1 | US-10-114-104-106 | Sequence 106, App |
| 15 | 221.5 | 9.1 | 263 | 1 | US-09-864-761-37656 | Sequence 37656, A |
| 16 | 111 | 4.5 | 120 | 1 | US-10-133-036-4 | Sequence 4, Appli |

ALIGNMENTS

RESULT 1

US-11-028-539-35
; Sequence 35, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (124)..(124)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (669)..(669)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (685)..(685)
; OTHER INFORMATION: Xaa is any amino acid
US-11-028-539-35

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.0541 | Length: | 685 |
| Score: | 641.00 | Matches: | 654 |
| Percent Similarity: | 95.5% | Conservative: | 0 |
| Best Local Similarity: | 95.5% | Mismatches: | 30 |
| Query Match: | 26.2% | Indels: | 1 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-11-028-539-35 (1-685)

```
Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ProLysThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeu 20

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
             ||||||||||||||||||||||||   ||||||||||||||||||||||||||||   |||
Db         21 LysThrLeuGlnGlyThrTyrPro***GluGluGlyLysGluLeuPheHisProCysAsp 40

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
             ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         41 MetValLeuValLysSerLeuProSerAsnSerProSerLeuAspThrSerTrpGluGly 60

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db         61 ProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyValGluSerTrp 80
```

```
Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
Db         81 IleHisHisThr***ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAsp 100

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
Db        101 AsnAlaSerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnPro 120

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
Db        121 GlyGlyLys***LeuLysSer***Ile-ProMetAlaLeuProTyrHisIlePheLeuPh 140

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
Db        140 eThrValLeuLeuProSerPheThrLeuThrAlaProProProCysArgCysMetThrSe 160

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
Db        160 rSerSerProTyrGlnGluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaPr 180

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
Db        180 oSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMetProAr 200

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
Db        200 gAsnCysTyrHisSerAlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLy 220

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
Db        220 sMetIleAsnProSerCysProGlyGlyLeuGlyValThrValCysTrpThrTyrPheTh 240

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
Db        240 rGlnThrGlyMetSerAspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLy 260

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
Db        260 sGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLysGlyLeuAs 280

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
Db        280 pLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnTh 300

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
Db        300 rThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLe 320

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
Db        320 uProLeuAsnPheArgProTyrValSerIleProValProGluGlnTrpAsnAsnPheSe 340

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
Db        340 rThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsnLeuGluIleTh 360

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
Db        360 rHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGl 380

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
Db        380 nCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePh 400

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
Db        400 eValCysGlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLe 420

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
Db        420 uSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIl 440

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
Db        440 eSerLysProArgAsnLysArgValProIleLeuProPheValIleGlyAlaGlyValLe 460

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
Db        460 uGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyrTyrLysLe 480

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
Db        480 uSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuValThrLeuGlnAs 500

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
Db        500 pGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuTh 520

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
Db        520 rAlaGluArgGlyGlyThrCysLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGl 540

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
Db        540 nSerGlyIleValThrGluLysValLysGluIleArgAspArgIleGlnArgArgAlaGl 560

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
Db        560 uGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuPr 580

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
Db        580 oPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAs 600

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
Db        600 nLeuLeuValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLy 620

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
Db        620 sMetGlnSerLysThrLysIleTyrArgArgProLeuAspArgProAlaSerProArgSe 640

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
Db        640 rAspValAsnAspIleLysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLe 660

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
Db        660 uArgProAsnSerAlaGlySerSer***SerGlyArgArgProThrSerProThrAlaLe 680

Qy       7230 TAGGTTTTCCTGT 7242
Db        680 uArgPheSerCys 684
```

RESULT 2
US-10-821-234-1395
; Sequence 1395, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1395
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-821-234-1395

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.265 | Length: | 540 |
| Score: | 518.00 | Matches: | 518 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.2% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-10-821-234-1395 (1-540)

```
Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
             |||||||  ||||||||||||||||||||||      |||      |||   |||||||||
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
             |||||||||||||||||||||||||||||||||||||||||||||||      ||||||||||||
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
             |||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
             |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
             |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
             |||  ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
             ||||||  ||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
             |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
             |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
             |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540
```

RESULT 3
US-11-028-539-26
; Sequence 26, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUM
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-028-539-26

Alignment Scores:

```
Pred. No.:              0.265          Length:         540
Score:                  518.00         Matches:        518
Percent Similarity:     96.1%          Conservative:   0
Best Local Similarity:  96.1%          Mismatches:     21
Query Match:            21.2%          Indels:         0
DB:                     1              Gaps:           0

US-10-717-580-11-COPY (1-7582) x US-11-028-539-26 (1-540)

Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
              ||||||  ||||||||||||||||||||||      |||     |||   |||||||||
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
              |||||||||||||||||||||||||||||||||||||||      ||||||||||||
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
              ||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
              |||   ||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
              ||||||   ||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540

RESULT 4
US-10-133-036-2
; Sequence 2, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
;  APPLICANT: Conrad, Bernard
;  APPLICANT:  Mach, Bernard
;  TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
;  FILE REFERENCE: 23135-507
;  CURRENT APPLICATION NUMBER: US/10/133,036
;  CURRENT FILING DATE:  2002-04-26
;  PRIOR APPLICATION NUMBER: PCT/EP00/10659
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 538
;   TYPE: PRT
;   ORGANISM: Human endogenous retrovirus
US-10-133-036-2

Alignment Scores:
Pred. No.:              0.269          Length:         538
Score:                  517.00         Matches:        517
Percent Similarity:     96.1%          Conservative:   0
Best Local Similarity:  96.1%          Mismatches:     21
Query Match:            21.1%          Indels:         0
DB:                     1              Gaps:           0
```

US-10-717-580-11-COPY (1-7582) x US-10-133-036-2 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 5
US-10-133-036-20
; Sequence 20, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Retrovirus
US-10-133-036-20

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.269 | Length: | 538 |
| Score: | 517.00 | Matches: | 517 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-10-133-036-20 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetGlyLeuProTyrHisIlePheLeuCysSerValLeuSerProCysPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
```

```
Db         21 AlaProProProCysArgCysMetThrSerSerSerProHisProGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlyValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnValGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780

Db        381 ArgAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValGluGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleProAspArgIleGlnArgIleAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerArgTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAspLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 6
US-10-133-036-21
; Sequence 21, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
;  APPLICANT: Conrad, Bernard
;  APPLICANT:  Mach, Bernard
;  TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
;  FILE REFERENCE: 23135-507
;  CURRENT APPLICATION NUMBER: US/10/133,036
;  CURRENT FILING DATE:  2002-04-26
;  PRIOR APPLICATION NUMBER: PCT/EP00/10659
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 538
;   TYPE: PRT
;   ORGANISM: Retrovirus
US-10-133-036-21

Alignment Scores:
Pred. No.:              0.269          Length:          538
Score:                  517.00         Matches:         517
Percent Similarity:     96.1%          Conservative:    0
Best Local Similarity:  96.1%          Mismatches:      21
Query Match:            21.1%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-10-133-036-21 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValValSerProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuCysLysGlyThrPro 60
```

```
Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet  80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetAlaIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 ArgAsnArgArgAlaLeuAspSerLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 7
US-10-717-580-33
; Sequence 33, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEA
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Human
US-10-717-580-33

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.269 | Length: | 538 |
| Score: | 517.00 | Matches: | 517 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-10-717-580-33 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetGlyLeuProTyrHisIlePheLeuCysSerValLeuSerProCysPheThrLeuThr  20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProHisProGluPheLeuTrpArg  40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
```

```
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrGlyValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnValGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ArgAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValGluGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||  ||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        421 IleProAspArgIleGlnArgIleAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerArgTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAspLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 8
US-09-902-535-2
; Sequence 2, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
;  APPLICANT: Keith, Jr., James C.
;  APPLICANT:  McCoy, John M.
;  APPLICANT:  Mi, Sha
;  TITLE OF INVENTION: Methods and compositions for diagnosing
;  TITLE OF INVENTION:  and treating preeclampsia and gestational trophoblast
;  TITLE OF INVENTION:  disorders
;  FILE REFERENCE: GIN-6006B4
;  CURRENT APPLICATION NUMBER: US/09/902,535
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: 60/216,657
;  PRIOR FILING DATE: 2000-07-06
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 538
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-535-2

Alignment Scores:
Pred. No.:              0.272          Length:          538
Score:                  516.00         Matches:         516
Percent Similarity:     95.9%          Conservative:    0
Best Local Similarity:  95.9%          Mismatches:      22
Query Match:            21.1%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-09-902-535-2 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||  ||||||||||||||||||||      |||      |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538

RESULT 9
US-10-114-893-135
; Sequence 135, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT: McCoy, John M.
;  APPLICANT: LaVallie, Edward R.
;  APPLICANT: Collins-Racie, Lisa A.
;  APPLICANT: Evans, Cheryl
;  APPLICANT: Merberg, David
;  APPLICANT: Treacy, Maurice
;  APPLICANT: Bowman, Michael R.
;  APPLICANT: Spaulding, Vikki
;  APPLICANT: Carlin-Duckett, McKeough
;  APPLICANT: Kelleher, Kerry S.
;  APPLICANT: Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: GI 6000-10A
;  CURRENT APPLICATION NUMBER: US/10/114,893
;  CURRENT FILING DATE: 2002-04-02
;  EARLIER APPLICATION NUMBER: 09/413,232
;  EARLIER FILING DATE: 1999-10-06
;  NUMBER OF SEQ ID NOS: 321
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
;   LENGTH: 538
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-114-893-135

Alignment Scores:
Pred. No.:                0.272          Length:          538
Score:                    516.00         Matches:         516
Percent Similarity:       95.9%          Conservative:    0
Best Local Similarity:    95.9%          Mismatches:      22
Query Match:              21.1%          Indels:          0
DB:                       1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-10-114-893-135 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||  ||||||||||||||||||||      |||      |||  ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60
```

```
Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet  80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 10
US-10-016-249-4
; Sequence 4, Application US/10016249
; Publication No. US20030100053A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-249-4

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.272 | Length: | 538 |
| Score: | 516.00 | Matches: | 516 |
| Percent Similarity: | 95.9% | Conservative: | 0 |
| Best Local Similarity: | 95.9% | Mismatches: | 22 |
| Query Match: | 21.1% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-10-016-249-4 (1-538)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr  20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg  40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro  60
```

```
Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet  80

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy      7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy      7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 11
US-10-416-642-2
; Sequence 2, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CD1
US-10-416-642-2

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.565 | Length: | 542 |
| Score: | 448.00 | Matches: | 460 |
| Percent Similarity: | 84.9% | Conservative: | 0 |
| Best Local Similarity: | 84.9% | Mismatches: | 78 |
| Query Match: | 18.3% | Indels: | 4 |
| DB: | 1 | Gaps: | 1 |

US-10-717-580-11-COPY (1-7582) x US-10-416-642-2 (1-542)

```
Qy      5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             |||  |||||||||  |||||||||      |||      |||  |||  ||||||
Db          1 MetAlaLeuProTyrCysIlePheLeuPheThrValLeuSerProProPheSerLeuThr 20

Qy      5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||  ||||||  |||  ||||||||||||||      ||||||||||||
Db         21 AlaProSerProCysHisCysArgThrSerSerSerProTyrGlnAlaPheLeuTrpArg 40

Qy      5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             |||  ||||||  ||||||||||||||||||||||||||||  ||||||  |||
Db         41 MetArgArgProArgHisIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnPro 60

Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||  |||||||||  ||||||||||||||||||||
```

```
Db         61 AlaPheThrAlaHisThrHisMetProHisAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||   |||   ||||||||||||||||||   ||||||||||||||||||||||||||||
Db        101 GlyAlaThrIleCysTrpThrTyrPheThrHisThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
                 |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerThrProSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrHisLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              ||||||||||||||||||||   ||||||||||||   |||   ||||||   |||
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerVal 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerIleLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleIleHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||      ||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrIleAlaThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              ||||||||||||||||||      |||      ||||||   ||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              ||||||||||||||   ||||||||||||||   ||||||||||||||||||   |||
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                ||||||||||||||||||   |||||||||   ||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              |||   ||||||||||||      ||||||||||||   ||||||      |||||||||
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluHisTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              ||||||   ||||||||||||   ||||||||||||||||||||||||||||   ||||||
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||||   |||||||||   ||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAA-----------CTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              ||||||||||||           ||||||||||||||   |||||||||   ||||||
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMetGluProGlnMetGlnSerMetThrLys 500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              ||||||      ||||||||||||||   ||||||   ||||||||||||||||||
Db        501 IleTyrHisGlyProLeuAspArgProAlaSerProCysSerAspValAsnAspIleGlu 520

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
              ||||||||||||   |||||||||   ||||||||||||||   ||||||||||||||
Db        521 GlyThrProProLysGluIleSerThrAlaGlnProLeuLeuCysProAsnSerAlaGly 540

Qy       7189 AGCAGT 7194
              ||||||
Db        541 SerSer 542

RESULT 12
US-10-637-565-20
; Sequence 20, Application US/10637565
; Publication No. US20040043381A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  PERRON, Herve
;  APPLICANT:  KOMURIAN-PRADEL, Florence
;  TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES F
;  TITLE OF INVENTION:  METHODS FOR DETECTING MSRV-1 RETROVIRUS
;  FILE REFERENCE: 110257
;  CURRENT APPLICATION NUMBER: US/10/637,565
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: US/09/890,340
;  PRIOR FILING DATE: 2001-11-27
;  PRIOR APPLICATION NUMBER: PCT/IB00/00159
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: EP 99420041.8
;  PRIOR FILING DATE: 1999-02-15
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 503
;   TYPE: PRT
;   ORGANISM: MSRV-1 retrovirus
US-10-637-565-20

Alignment Scores:
Pred. No.:                0.843          Length:          503
Score:                    418.00         Matches:         430
Percent Similarity:       85.5%          Conservative:    0
Best Local Similarity:    85.5%          Mismatches:      69
Query Match:              17.1%          Indels:          4
DB:                       1              Gaps:            1

US-10-717-580-11-COPY (1-7582) x US-10-637-565-20 (1-503)

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
              |||         ||||||||||||||||||||||||||||||||||   ||||||
Db          1 ArgThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsn 20

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
                 ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         21 SerThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCys 40

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 60
```

```
Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
Db         61 LeuGlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGly 80

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
Db         81 IleGlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGly 100

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
Db        101 HisSerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArg 120

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
Db        121 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSer 140

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
Db        141 AlaGlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSer 160

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
Db        161 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 180

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
Db        181 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 200

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
Db        201 PheSerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThr 220

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
Db        221 ArgIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCys 240

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
Db        241 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 260

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        261 TyrThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValPro 280

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
Db        281 IleLeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySer 300

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
Db        301 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGlu 320

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
Db        321 GlnValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 340

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
Db        341 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPhe 360

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
Db        361 LeuGlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLys 380

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
Db        381 GluIleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGly 400

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
Db        401 LeuLeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIle 420

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
Db        421 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArg 440

Qy       7018 ATCGAAGCTGTRAAA----------CTACAAATGGAGCCCAAGATGCAGTCCAAGACT 7065
Db        441 IleGluAlaValLysLeuGlnMetValLeuGlnMetGluProGlnMetGluSerMetThr 460

Qy       7066 AAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATC 7125
Db        461 LysIleHisArgGlyProLeuAspArgProAlaSerProCysSerAspValAsnAspIle 480

Qy       7126 AAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCA 7185
Db        481 GluGlyThrProProGluGluIleSerThrAlaGlnProLeuLeuCysProAsnSerAla 500

Qy       7186 GGAAGCAGT 7194
Db        501 GlySerSer 503
```

RESULT 13
US-08-979-847-106
; Sequence 106, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-979-847-106

Alignment Scores:
Pred. No.: 0.879 Length: 493
Score: 416.00 Matches: 416

```
Percent Similarity:     84.7%          Conservative:  0
Best Local Similarity:  84.7%          Mismatches:    75
Query Match:            17.0%          Indels:        0
DB:                     1              Gaps:          0

US-10-717-580-11-COPY (1-7582) x US-08-979-847-106 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 14
US-10-114-104-106
; Sequence 106, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
```

```
;              REGISTRATION NUMBER: 30,024
;              REFERENCE/DOCKET NUMBER: WPB 39046A
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-836-6400
;              TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 106:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 493 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-10-114-104-106

Alignment Scores:
Pred. No.:                0.879          Length:          493
Score:                    416.00         Matches:         416
Percent Similarity:       84.7%          Conservative:    0
Best Local Similarity:    84.7%          Mismatches:      75
Query Match:              17.0%          Indels:          0
DB:                       1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-10-114-104-106 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 15
US-09-864-761-37656
; Sequence 37656, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
```

```
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37656
;   LENGTH: 263
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC002346.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  20
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
;   OTHER INFORMATION: EST_HUMAN HIT: AU138405.1, EVALUE 7.00e-90
;   OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 1.00e-16
US-09-864-761-37656

Alignment Scores:
Pred. No.:                12.6          Length:          263
Score:                    221.50        Matches:         232
Percent Similarity:       87.9%         Conservative:    0
Best Local Similarity:    87.9%         Mismatches:      31
Query Match:              9.1%          Indels:          1
DB:                       1             Gaps:            1

US-10-717-580-11-COPY (1-7582) x US-09-864-761-37656 (1-263)

Qy       5956 GAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCC 6015
              |||||||||   |||||||||||||||||||||||||   ||||||   |||       |||
Db          1 GluLysHisIleLysGluValIleSerGlnLeuThrTrpValHisSerThrProGlyPro 20

Qy       6016 TACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTA 6075
              |||||||||||||||||||||||||||||||||||   |||||||||   ||||||
Db         21 TyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuHisThrHisThrGlyLeuVal 40

Qy       6076 AGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAC 6135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsn 60

Qy       6136 TGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAA 6195
              ||||||   ||||||||||||   |||   ||||||   |||||||||   |||||||||
Db         61 CysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIleProIleProGluGln 80

Qy       6196 TGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCC 6255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         81 TrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyProLeu---Ser 99

Qy       6256 AATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrIleAsp 119

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              |||   ||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db        120 ThrAlaAsnSerGlnCysIleArgTrpValThrProProThrArgIleValCysLeuPro 139

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        140 SerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeuAsnGlySerSerGlu 159

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              |||   ||||||||||||||||||||   ||||||   ||||||||||||||||||||
Db        160 SerValCysPheLeuSerPheLeuValAlaProMetProIleTyrThrGluGlnAspLeu 179

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        180 TyrAsnHisValIleProLysProArgAsnLysArgValProIleLeuProPheValIle 199

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              ||||||||||||||||          ||||||||||||||||||||||||||||||||||
Db        200 GlyAlaGlyValLeuGlyGlyValAlaThrGlyIleGlyGlyIleThrThrSerThrGln 219

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              ||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        220 PheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluTrpValAlaAspThrLeu 239

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        240 ValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAla 259

Qy       6736 TTAGACTYGCTA 6747
              ||||||   |||
Db        260 LeuAspLeuLeu 263

RESULT 16
US-10-133-036-4
; Sequence 4, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
;  APPLICANT: Conrad, Bernard
;  APPLICANT:  Mach, Bernard
;  TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
;  FILE REFERENCE: 23135-507
;  CURRENT APPLICATION NUMBER: US/10/133,036
;  CURRENT FILING DATE:  2002-04-26
;  PRIOR APPLICATION NUMBER: PCT/EP00/10659
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 120
;   TYPE: PRT
;   ORGANISM: Human endogenous retrovirus
US-10-133-036-4

Alignment Scores:
Pred. No.:                72.3          Length:          120
Score:                    111.00        Matches:         111
Percent Similarity:       92.5%         Conservative:    0
Best Local Similarity:    92.5%         Mismatches:      9
Query Match:              4.5%          Indels:          0
DB:                       1             Gaps:            0
```

US-10-717-580-11-COPY (1-7582) x US-10-133-036-4 (1-120)

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   |||||||||||||||||||||      |||     |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120
```

Search completed: August 4, 2006, 17:09:52
Job time : 31 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 4, 2006, 16:51:17 ; Search time 1.1 Seconds
(without alignments)
6.303 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 2447
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: UNITARY_AA
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Issued Patents_AA

Searched: 9 seqs, 3048 residues

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 9 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=new.seq -DB=aaissdb -SUFFIX=pto
-OUT=aaiss.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=unitary_AA -TRANS=human40.cdi -LIST=9 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=9 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : aaissdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 641 | 26.2 | 685 | 1 | US-09-719-554-35 | Sequence 35, Appl |
| 2 | 518 | 21.2 | 540 | 1 | US-09-949-016-11511 | Sequence 11511, A |
| 3 | 518 | 21.2 | 540 | 1 | US-09-719-554-26 | Sequence 26, Appl |
| 4 | 516 | 21.1 | 538 | 1 | US-09-175-928-4 | Sequence 4, Appli |
| 5 | 416 | 17.0 | 493 | 1 | US-08-979-847B-106 | Sequence 106, App |
| c 6 | 69 | 2.8 | 685 | 1 | US-09-719-554-35 | Sequence 35, Appl |
| 7 | 62 | 2.5 | 63 | 1 | US-09-719-554-30 | Sequence 30, Appl |
| 8 | 58 | 2.4 | 67 | 1 | US-08-686-878A-49 | Sequence 49, Appl |
| 9 | 58 | 2.4 | 67 | 1 | US-08-721-489-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-09-719-554-35
; Sequence 35, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (124)..(124)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (669)..(669)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (685)..(685)
; OTHER INFORMATION: Xaa is any amino acid
US-09-719-554-35

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.00863 | Length: | 685 |
| Score: | 641.00 | Matches: | 654 |
| Percent Similarity: | 95.5% | Conservative: | 0 |
| Best Local Similarity: | 95.5% | Mismatches: | 30 |
| Query Match: | 26.2% | Indels: | 1 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-09-719-554-35 (1-685)

```
Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ProLysThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeu 20

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
                |||||||||||||||||||||||  |||||||||||||||||||||||||||||   |||
Db         21 LysThrLeuGlnGlyThrTyrPro***GluGluGlyLysGluLeuPheHisProCysAsp 40

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
                |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db         41 MetValLeuValLysSerLeuProSerAsnSerProSerLeuAspThrSerTrpGluGly 60

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
                ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db         61 ProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyValGluSerTrp 80

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
                ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleHisHisThr***ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAsp 100

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AsnAlaSerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnPro 120

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
                |||||||||  |||||||||  ||| ||||||  ||||||||||||||||||||||
Db        121 GlyGlyLys***LeuLysSer***Ile-ProMetAlaLeuProTyrHisIlePheLeuPh 140

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
```

```
Db        140 eThrValLeuLeuProSerPheThrLeuThrAlaProProProCysArgCysMetThrSe 160

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
Db        160 rSerSerProTyrGlnGluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaPr 180

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
Db        180 oSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMetProAr 200

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
Db        200 gAsnCysTyrHisSerAlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLy 220

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
Db        220 sMetIleAsnProSerCysProGlyGlyLeuGlyValThrValCysTrpThrTyrPheTh 240

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
Db        240 rGlnThrGlyMetSerAspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLy 260

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
Db        260 sGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLysGlyLeuAs 280

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
Db        280 pLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnTh 300

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
Db        300 rThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLe 320

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
Db        320 uProLeuAsnPheArgProTyrValSerIleProValProGluGlnTrpAsnAsnPheSe 340

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
Db        340 rThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsnLeuGluIleTh 360

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
Db        360 rHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGl 380

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
Db        380 nCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePh 400

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
Db        400 eValCysGlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLe 420

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
Db        420 uSerPheLeuValProProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIl 440

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
Db        440 eSerLysProArgAsnLysArgValProIleLeuProPheValIleGlyAlaGlyValLe 460

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
Db        460 uGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyrTyrLysLe 480

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
Db        480 uSerGlnGluLeuAsnGlyAspMetGluArgValAlaAspSerLeuValThrLeuGlnAs 500

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
```

```
Db        500 pGlnLeuAsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuTh 520

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
Db        520 rAlaGluArgGlyGlyThrCysLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGl 540

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
Db        540 nSerGlyIleValThrGluLysValLysGluIleArgAspArgIleGlnArgArgAlaGl 560

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
Db        560 uGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuPr 580

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
Db        580 oPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAs 600

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
Db        600 nLeuLeuValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLy 620

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
Db        620 sMetGlnSerLysThrLysIleTyrArgArgProLeuAspArgProAlaSerProArgSe 640

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
Db        640 rAspValAsnAspIleLysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLe 660

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
Db        660 uArgProAsnSerAlaGlySerSer***SerGlyArgArgProThrSerProThrAlaLe 680

Qy       7230 TAGGTTTTCCTGT 7242
Db        680 uArgPheSerCys 684

RESULT 2
US-09-949-016-11511
; Sequence 11511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11511
;   LENGTH: 540
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11511

Alignment Scores:
Pred. No.:                0.0495          Length:          540
Score:                    518.00          Matches:         518
Percent Similarity:       96.1%           Conservative:    0
Best Local Similarity:    96.1%           Mismatches:      21
Query Match:              21.2%           Indels:          0
DB:                       1               Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-09-949-016-11511 (1-540)
```

```
Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
             ||||||  ||||||||||||||||||||      |||      |||  |||||||||
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
             ||||||||||||||||||||||||||||||||||||||||||||      ||||||||||||
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
             ||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
             ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
             ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777
             |||  ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
             ||||||  ||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
             |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
             ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540
```

RESULT 3
US-09-719-554-26
; Sequence 26, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUI
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-719-554-26

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0.0495 | Length: | 540 |
| Score: | 518.00 | Matches: | 518 |
| Percent Similarity: | 96.1% | Conservative: | 0 |
| Best Local Similarity: | 96.1% | Mismatches: | 21 |
| Query Match: | 21.2% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-717-580-11-COPY (1-7582) x US-09-719-554-26 (1-540)

```
Qy       5578 CCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTC 5637
             ||||||  ||||||||||||||||||||      |||      |||  |||||||||
Db          2 ProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeu 21

Qy       5638 ACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGG 5697
```

```
Db         22 ThrAlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrp 41

Qy       5698 AGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACC 5757
Db         42 ArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThr 61

Qy       5758 CCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGC 5817
Db         62 ProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCys 81

Qy       5818 ATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGA 5877
Db         82 MetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGly 101

Qy       5878 CTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGA 5937
Db        102 LeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGly 121

Qy       5938 GTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTA 5997
Db        122 ValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgVal 141

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
Db        142 HisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArg 161

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
Db        162 ThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSer 181

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
Db        182 AlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSer 201

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
Db        202 IleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeu 221

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
Db        222 ValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLys 241

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
Db        242 PheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThr 261

Qy       6358 CAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGT 6417
Db        262 GlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCys 281

Qy       6418 TTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATC 6477
Db        282 LeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIle 301

Qy       6478 TACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
Db        302 TyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValPro 321

Qy       6538 ATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT 6597
Db        322 IleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341

Qy       6598 ATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAA 6657
Db        342 IleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGlu 361

Qy       6658 CGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTC 6717
Db        362 ArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValVal 381

Qy       6718 CTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTT 6777

Db        382 LeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPhe 401

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
Db        402 LeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLys 421

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
Db        422 GluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGly 441

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
Db        442 LeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIle 461

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
Db        462 LeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArg 481

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
Db        482 IleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArg 501

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCT 7137
Db        502 ArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrPro 521

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
Db        522 ProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 540

RESULT 4
US-09-175-928-4
; Sequence 4, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT: McCoy, John M.
;  APPLICANT: LaVallie, Edward R.
;  APPLICANT: Collins-Racie, Lisa A.
;  APPLICANT: Evans, Cheryl
;  APPLICANT: Merberg, David
;  APPLICANT: Treacy, Maurice
;  APPLICANT: Mi, Sha
;  APPLICANT: Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: 6006B.AJ172A
;  CURRENT APPLICATION NUMBER: US/09/175,928A
;  CURRENT FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 538
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-175-928-4

Alignment Scores:
Pred. No.:              0.0509          Length:         538
Score:                  516.00          Matches:        516
Percent Similarity:     95.9%           Conservative:   0
Best Local Similarity:  95.9%           Mismatches:     22
Query Match:            21.1%           Indels:         0
DB:                     1               Gaps:           0

US-10-717-580-11-COPY (1-7582) x US-09-175-928-4 (1-538)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40
```

```
Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSerAspGlyGlyGlyVal 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        121 GlnAspGlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        181 GlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyrValSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpValThrProProThrGln 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrArgCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        301 ThrGluGlnAspLeuTyrAsnTyrValIleSerLysProArgAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsnGlyAspMetGluArg 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
                 ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        401 GlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             |||   ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        421 IleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeuAlaAlaIleIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThrLysIleTyrArgArg 500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProPro 520

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer 538
```

RESULT 5
US-08-979-847B-106
; Sequence 106, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-08-979-847B-106

Alignment Scores:
Pred. No.:              0.207          Length:          493
Score:                  416.00         Matches:         416
Percent Similarity:     84.7%          Conservative:    0
Best Local Similarity:  84.7%          Mismatches:      75
Query Match:            17.0%          Indels:          0
DB:                     1              Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-08-979-847B-106 (1-493)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 MetAlaLeuProTyrHisThrPheLeuPheThrValLeuLeuProProPheAlaLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         21 AlaProProProCysCysCysThrThrSerSerSerProTyrGlnGluPheLeu***Arg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db         41 ThrArgLeuProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyAsnSer 60

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db         61 ThrPheThrAlaHisThrHisMetProArgAsnCysTyrAsnSerAlaThrLeuCysMet 80

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         81 HisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSerCysProGlyGlyLeu 100

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        101 GlyAlaThrValCysTrpThrTyrPheThrHisThrSerMetSerAspGlyGlyGlyIle 120

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        121 GlnGlyGlnAlaArgGluLysGlnValLysGluAlaIleSerGlnLeuThrArgGlyHis 140

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        141 SerThrProSerProTyrLysGlyLeuValLeuSerLysLeuHisGluThrLeuArgThr 160

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        161 HisThrArgLeuValSerLeuPheAsnThrThrLeuThrArgLeuHisGluValSerAla 180

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        181 GlnAsnProThrAsnCysTrpMetCysLeuProLeuHisPheArgProTyrIleSerIle 200

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        201 ProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuVal 220

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        221 GlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPhe 240

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        241 SerAsnThrIleAspThrThrSerSerGlnCysIleArgTrpValThrProProThrArg 260

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        261 IleValCysLeuProSerGlyIlePhePheValCysGlyThrSerAlaTyrHisCysLeu 280

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        281 AsnGlySerSerGluSerMetCysPheLeuSerPheLeuValProProMetThrIleTyr 300

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        301 ThrGluGlnAspLeuTyrAsnHisValValProLysProHisAsnLysArgValProIle 320

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        321 LeuProPheValIleArgAlaGlyValLeuGlyArgLeuGlyThrGlyIleGlySerIle 340

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        341 ThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluIleAsnGlyAspMetGluGln 360

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        361 ValThrAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeuAlaAlaValValLeu 380

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        381 GlnAsnArgArgAlaLeuAspLeuLeuThrAlaLysArgGlyGlyThrCysLeuPheLeu 400

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        401 GlyGluGluArgCysTyrTyrValAsnGlnSerArgIleValThrGluLysValLysGlu 420

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db        421 IleArgAspArgIleGlnCysArgAlaGluGluLeuGlnAsnThrGluArgTrpGlyLeu 440

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        441 LeuSerGlnTrpMetProTrpValLeuProPheLeuGlyProLeuAlaAlaLeuIleLeu 460

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        461 LeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValLysPheValSerSerArgIle 480

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
Db        481 GluAlaValLysLeuGlnMetValLeuGlnMet 491

RESULT 6
US-09-719-554-35
; Sequence 35, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUM
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 685
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(29)
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: (85)..(85)
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: (124)..(124)
```

```
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: (128)..(128)
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: (669)..(669)
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: Xaa is any amino acid
US-09-719-554-35

Alignment Scores:
Pred. No.:                8.47          Length:          685
Score:                    69.00         Matches:         69
Percent Similarity:       10.2%         Conservative:    0
Best Local Similarity:    10.2%         Mismatches:      608
Query Match:              2.8%          Indels:          0
DB:                       1             Gaps:            0

US-10-717-580-11-COPY (1-7582) x US-09-719-554-35 (1-685)

Qy       7568 TTAATAGAGTGAAATAGCATGAAAACAGAGCTCCCATACAAAGGGAGGGGACCCAAAGGG 7509
                 |||                                                   |||
Db          6 LeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGly 25

Qy       7508 GGTTGCCGTTGCCGGCTCGAARACTTGGGTTTATATCCCGATCATTGTCCCTCCTGCTGT 7449
                                                    |||
Db         26 ThrTyrPro***GluGluGlyLysGluLeuPheHisProCysAspMetValLeuValLys 45

Qy       7448 GCTCTCAKGCAATARATGATTGGCTATTTCTTTACCTCCTGTCTTTGCCTAATTAGCATT 7389
                  |||                                                   |||
Db         46 SerLeuProSerAsnSerProSerLeuAspThrSerTrpGluGlyProTyrProValIle 65

Qy       7388 TTAGTGAGCTCTCTGATTGGTCAGGTGTGARCTAAGTTGCAAGCCCCGTGTTTAAAGGTG 7329
                 |||                     |||
Db         66 LeuSerThrProThrAlaValLysValAlaGlyValGluSerTrpIleHisHisThr*** 85

Qy       7328 GATGTGGTCACCTTCCCASCTAGGCTTARGGATTCTTARTCAGCCTAGGAAATCCAGCTA 7269

Db         86 ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAlaSerTyrSer 105

Qy       7268 GTCCTGTCTCTCAGTCCCCCATCTCAACAGGAAAACCTAAGTGCTGTTGGGGAGGTTGGC 7209
                    |||                                             |||
Db        106 CysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys***Leu 125

Qy       7208 CGACSACCGCTCTAACTGCTTCCTGCTGAATTGGGGCGTAGTAGAGGTTGTGCAGCTGAG 7149

Db        126 LysSer***IleProMetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuPro 145

Qy       7148 ATTTCCTCAGGAGGGGTGCCTTTGATGTCATTAACATCAGATCGTGGGCTARCAGGCCGG 7089
                              |||
Db        146 SerPheThrLeuThrAlaProProProCysArgCysMetThrSerSerSerProTyrGln 165

Qy       7088 TCCAGGGGTCTGCGGTAGATCTTAGTCTTGGACTGCATCTTGGGCTCCATTTGTAGTTTY 7029
                    |||                                     |||     |||
Db        166 GluPheLeuTrpArgMetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeu 185

Qy       7028 ACAGCTTCGATTCTGGAAGAGACAAAGTTAACAAGGAGGTYAAAGATACAGGGTCCAAAG 6969
                              |||     |||
Db        186 SerLysGlyThrProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSer 205

Qy       6968 AGGAGTAGCAATATTATAGCTGCTAGAGGTCCTAAGAAGGGGAGAATCCAGGGCATCCAT 6909
                              |||                 |||
Db        206 AlaThrLeuCysMetHisAlaAsnThrHisTyrTrpThrGlyLysMetIleAsnProSer 225

Qy       6908 YGGCTGAGGAGGCCCCAGGGTCCAGTGTTTCGAAGCTCYTCTGCTMTACGTTGTATTCGA 6849
                                       |||
Db        226 CysProGlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMetSer 245

Qy       6848 TCTSGAATTTCTTYAACTTTCTCAGTGACGATTCCGGATTGATTAACATAATAACAGCAT 6789

Db        246 AspGlyGlyGlyValGlnAspGlnAlaArgGluLysHisValLysGluValIleSerGln 265

Qy       6788 TCTTCCCCTAAAAATAAACAGGTTCCCCCTCTYTCAGCGGTTAGCRAGTCTAAAGCTCTT 6729
                                          |||                  ||||||
Db        266 LeuThrArgValHisGlyThrSerSerProTyrLysGlyLeuAspLeuSerLysLeuHis 285

Qy       6728 CGATTTYGAAGGACTACTGCTGCTAGGGAGTTAAGTTGATCTTGCAAGGTGACCAGGGAG 6669
                    ||||||                                          |||
Db        286 GluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeuThrGlyLeu 305

Qy       6668 TCGGCGACCCGTTCCATGTCCCCATTTAGTTCTTGAGATAGTTTGTAGTAGAACTGAGTA 6609
                                       |||                 |||     |||
Db        306 HisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArg 325

Qy       6608 GAGGTTGTGATACCGCCAATGCCAGTACCTAGTGCACCTAGCACTCCTGCTCCTATAACA 6549
                  |||     |||   |||                                   |||
Db        326 ProTyrValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThr 345

Qy       6548 AAAGGAAGAATGGGTACTCTTTTGTTGCGGGGCTTAGATATGACATAACTGTATAAATCT 6489
                                       |||         |||     |||
Db        346 ThrSerValLeuValGlyProLeuValSerAsnLeuGluIleThrHisThrSerAsnLeu 365

Qy       6488 TGTTCAGTGTAGATGGYCATRGGGGGCACTAAGAATGAGAGGAAGCACATAGATTCTGAA 6429
                  |||                                             |||
Db        366 ThrCysValLysPheSerAsnThrThrTyrThrThrAsnSerGlnCysIleArgTrpVal 385

Qy       6428 GAGCCATTCAAACAACGATAGGCTGAGGTACCACAGACAAAAAATATTCCTGAGGGTAGG 6369
                    |||     |||                                         |||
Db        386 ThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCysGlyThrSer 405

Qy       6368 CAGACTATTTGTGTGGGAGGAGTTACCCACCTGATGCATTGGGAGTTGGTTGTGTATGTA 6309
                    |||     |||         |||
Db        406 AlaTyrArgCysLeuAsnGlySerSerGluSerMetCysPheLeuSerPheLeuValPro 425

Qy       6308 GTATTGCTAAATTTTACACAGGTGAGGTTTGAGGTATGGGTTATTTCCASATTGGAAACA 6249
                                       |||                 |||||||||
Db        426 ProMetThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsn 445

Qy       6248 AGAGGTCCTACTAAAACGGAAGTGGTGTTTATTTCTGTGCTGAAGTTGTTCCATTGTTCA 6189
                                          |||
Db        446 LysArgValProIleLeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThr 465

Qy       6188 GGTACAGGGATTGAAACATATGGCYTGAAGTTCAGGGGGAGGCATATCCAACAGTTAGTA 6129
                  |||   |||     |||         |||             |||   |||
Db        466 GlyIleGlyGlyIleThrThrSerThrGlnPheTyrTyrLysLeuSerGlnGluLeuAsn 485

Qy       6128 GGGTTTTGGGCCGAGACCTCATGGAGCCCAGTGAGGGTGGTATTAAATAGGCTTACCAGG 6069
                  |||                 |||   |||
Db        486 GlyAspMetGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeu 505

Qy       6068 CGAGTATGGGTACGGAGGGTTTCATGTAGTTTTGAGAGATCTAGTCCTTTGTAGGGGCTA 6009
                       |||                                         |||
Db        506 AlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGly 525

Qy       6008 GAGGTGCCATGTACCCSGGTGAGTTGGGAGATTACTTCTTTTACATGTTTTTCTCTTGCC 5949

Db        526 ThrCysLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThr 545

Qy       5948 TGATCTTGAACTCCACCCCCATCAGACATACCAGTTTGGGTGAAGTAAGTCCAACAGACA 5889
                                          |||                 |||
Db        546 GluLysValLysGluIleArgAspArgIleGlnArgArgAlaGluGluLeuArgAsnThr 565

Qy       5888 GTGACTCCAAGTCCTCCAGGACAACTAGGATTAATCATTTTTCCTGTCCAATAATGAGTA 5829
                                       |||         |||   |||
Db        566 GlyProTrpGlyLeuLeuSerGlnTrpMetProTrpIleLeuProPheLeuGlyProLeu 585

Qy       5828 TTTGCATGCATGCAAAGAGTGGCAGAGTGATAGCAGTTGCGGGGCATATGGGTGTGGGCA 5769
                  |||                                                 |||
Db        586 AlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeuValAsnPhe 605

Qy       5768 GTGAAGGTGGGGGTTCCCTTASAAAGACTCCTATACGATGGGGCATCAATATTTCCGGGA 5709
                  |||                 |||
Db        606 ValSerSerArgIleGluAlaValLysLeuGlnMetGluProLysMetGlnSerLysThr 625
```

```
Qy       5708 CGCTGCATTCTCCATAGAAACTCTKGGTRAGGGGAGCTACTGGTCATACAGCGGCATGGA 5649

Db        626 LysIleTyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIle 645

Qy       5648 GGGGGTGCAGTGAGAGTGAAASAGGGTRAAASAACASTAMAGAGAAAAATATGATAAGGG 5589
                 |||
Db        646 LysGlyThrProProGluGluIleSerAlaAlaGlnProLeuLeuArgProAsnSerAla 665

Qy       5588 AGGSCCATGGGGGATTTATGATTTTAGTTACTTTCCTCCTGGTTGTTGTTT 5538
                                                     |||   |||
Db        666 GlySerSer***SerGlyArgArgProThrSerProThrAlaLeuArgPhe 682

RESULT 7
US-09-719-554-30
; Sequence 30, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
;   LENGTH: 63
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-719-554-30

Alignment Scores:
Pred. No.:                29.3          Length:         63
Score:                    62.00         Matches:        62
Percent Similarity:       98.4%         Conservative:   0
Best Local Similarity:    98.4%         Mismatches:     1
Query Match:              2.5%          Indels:         0
DB:                       1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x US-09-719-554-30 (1-63)

Qy        834 GAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GluThrThrHisAsnTyrValLysSerValIleTyrAlaLeuGlnGluAlaPheArgVal 20

Qy        894 TACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACC 953
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         21 TyrLeuProIleProAlaSerProThrProSerProThrAsnLysAspProProSerThr 40

Qy        954 CAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATT 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlnMetValGlnLysGluIleAspLysArgValAsnSerGluProLysSerAlaAsnIle 60

Qy       1014 CCCCAATTA 1022
              |||||||||
Db         61 ProGlnLeu 63

RESULT 8
US-08-686-878A-49
; Sequence 49, Application US/08686878A
; Patent No. 5708157
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Evans, Cheryl
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/686,878A
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-686-878A-49

Alignment Scores:
Pred. No.:                30.2          Length:         67
Score:                    58.00         Matches:        58
Percent Similarity:       86.6%         Conservative:   0
Best Local Similarity:    86.6%         Mismatches:     9
Query Match:              2.4%          Indels:         0
DB:                       1             Gaps:           0

US-10-717-580-11-COPY (1-7582) x US-08-686-878A-49 (1-67)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   ||||||||||||||||||||||      |||      |||   ||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCAT 5781
              ||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHis 67

RESULT 9
US-08-721-489-3
; Sequence 3, Application US/08721489
; Patent No. 5786465
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
```

```
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/721,489
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-721-489-3

Alignment Scores:
Pred. No.:              30.2          Length:        67
Score:                  58.00         Matches:       58
Percent Similarity:     86.6%         Conservative:  0
Best Local Similarity:  86.6%         Mismatches:    9
Query Match:            2.4%          Indels:        0
DB:                     1             Gaps:          0

US-10-717-580-11-COPY (1-7582) x US-08-721-489-3 (1-67)

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              |||   ||||||||||||||||||||      |||      |||   ||||||||||||
Db          1 MetAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThr 20

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db         21 AlaProProProCysArgCysMetThrSerSerSerProTyrGlnGluPheLeuTrpArg 40

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         41 MetGlnArgProGlyAsnIleAspAlaProSerTyrArgSerLeuSerLysGlyThrPro 60

Qy       5761 ACCTTCACTGCCCACACCCAT 5781
              ||||||||||||||||||||
Db         61 ThrPheThrAlaHisThrHis 67


Search completed: August  4, 2006, 16:51:37
Job time : 13.5 secs
```

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2006, 13:09:06 ; Search time 661 Seconds
(without alignments)
4.600 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 7489.6
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 383 seqs, 200533 residues

Total number of hits satisfying chosen parameters: 766

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

EST

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 383 summaries

Database : estdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 2180.9 | 29.1 | 2748 | 1 | CR605851 | ACCESSION:CR605851 |
| | 2 | 2176.9 | 29.1 | 2749 | 1 | CR617248 | ACCESSION:CR617248 |
| | 3 | 2174.9 | 29.0 | 2500 | 1 | CR622175 | ACCESSION:CR622175 |
| | 4 | 2169.9 | 29.0 | 2748 | 1 | CR625046 | ACCESSION:CR625046 |
| | 5 | 2147.9 | 28.7 | 2716 | 1 | CR613169 | ACCESSION:CR613169 |
| c | 6 | 970 | 13.0 | 1071 | 1 | BX365066 | ACCESSION:BX365066 |
| c | 7 | 923.7 | 12.3 | 998 | 1 | BX337769 | ACCESSION:BX337769 |
| c | 8 | 912.6 | 12.2 | 1058 | 1 | BX378303 | ACCESSION:BX378303 |
| c | 9 | 909.1 | 12.1 | 935 | 1 | BX391741 | ACCESSION:BX391741 |
| | 10 | 887.2 | 11.8 | 955 | 1 | BX347952 | ACCESSION:BX347952 |
| c | 11 | 886 | 11.8 | 1019 | 1 | BX439636 | ACCESSION:BX439636 |
| | 12 | 867.4 | 11.6 | 1014 | 1 | BX378304 | ACCESSION:BX378304 |
| | 13 | 837.3 | 11.2 | 924 | 1 | BX347314 | ACCESSION:BX347314 |
| | 14 | 824.3 | 11.0 | 931 | 1 | BX389656 | ACCESSION:BX389656 |
| | 15 | 823 | 11.0 | 903 | 1 | BX408752 | ACCESSION:BX408752 |
| c | 16 | 813.5 | 10.9 | 872 | 1 | BX326647 | ACCESSION:BX326647 |
| | 17 | 807 | 10.8 | 883 | 1 | BX408733 | ACCESSION:BX408733 |
| c | 18 | 806.6 | 10.8 | 966 | 1 | BX380176 | ACCESSION:BX380176 |
| c | 19 | 791.2 | 10.6 | 877 | 1 | BX347619 | ACCESSION:BX347619 |
| | 20 | 779.701 | 10.4 | 921 | 1 | BX408734 | ACCESSION:BX408734 |
| | 21 | 774.602 | 10.3 | 995 | 1 | BX389657 | ACCESSION:BX389657 |
| c | 22 | 772 | 10.3 | 836 | 1 | BX388766 | ACCESSION:BX388766 |
| | 23 | 768.6 | 10.3 | 808 | 1 | DA834147 | ACCESSION:DA834147 |
| c | 24 | 766.4 | 10.2 | 844 | 1 | BX347111 | ACCESSION:BX347111 |
| | 25 | 765.8 | 10.2 | 828 | 1 | BX367907 | ACCESSION:BX367907 |
| c | 26 | 756.3 | 10.1 | 846 | 1 | BX368078 | ACCESSION:BX368078 |
| | 27 | 753.3 | 10.1 | 820 | 1 | DA833993 | ACCESSION:DA833993 |
| c | 28 | 752.6 | 10.0 | 900 | 1 | BX459153 | ACCESSION:BX459153 |
| | 29 | 734.7 | 9.8 | 836 | 1 | CZ457645 | ACCESSION:CZ457645 |
| | 30 | 733.6 | 9.8 | 770 | 1 | DR004257 | ACCESSION:DR004257 |
| | 31 | 731.6 | 9.8 | 769 | 1 | DA833923 | ACCESSION:DA833923 |
| | 32 | 729.9 | 9.7 | 879 | 1 | BX325902 | ACCESSION:BX325902 |
| c | 33 | 724.6 | 9.7 | 758 | 1 | BX357208 | ACCESSION:BX357208 |
| | 34 | 723.302 | 9.7 | 924 | 1 | BX409328 | ACCESSION:BX409328 |
| | 35 | 718.6 | 9.6 | 761 | 1 | DA832031 | ACCESSION:DA832031 |
| | 36 | 717 | 9.6 | 782 | 1 | DA833912 | ACCESSION:DA833912 |
| | 37 | 711.9 | 9.5 | 840 | 1 | DA834086 | ACCESSION:DA834086 |
| | 38 | 710.9 | 9.5 | 771 | 1 | BI087886 | ACCESSION:BI087886 |
| | 39 | 704.7 | 9.4 | 792 | 1 | DA830020 | ACCESSION:DA830020 |
| | 40 | 697.7 | 9.3 | 721 | 1 | AU138405 | ACCESSION:AU138405 |
| | 41 | 690.7 | 9.2 | 819 | 1 | DA830049 | ACCESSION:DA830049 |
| | 42 | 684 | 9.1 | 719 | 1 | CN272394 | ACCESSION:CN272394 |
| | 43 | 682.4 | 9.1 | 799 | 1 | CZ450650 | ACCESSION:CZ450650 |
| c | 44 | 682.2 | 9.1 | 832 | 1 | DU836611 | ACCESSION:DU836611 |
| | 45 | 677.9 | 9.1 | 702 | 1 | AU138097 | ACCESSION:AU138097 |
| | 46 | 676.1 | 9.0 | 795 | 1 | BZ611989 | ACCESSION:BZ611989 |
| | 47 | 667.6 | 8.9 | 714 | 1 | BE734284 | ACCESSION:BE734284 |
| | 48 | 665.3 | 8.9 | 775 | 1 | DA834377 | ACCESSION:DA834377 |
| | 49 | 657.5 | 8.8 | 850 | 1 | DA834266 | ACCESSION:DA834266 |
| | 50 | 650 | 8.7 | 717 | 1 | DA830985 | ACCESSION:DA830985 |
| | 51 | 646.5 | 8.6 | 698 | 1 | DA834382 | ACCESSION:DA834382 |
| | 52 | 645.2 | 8.6 | 769 | 1 | CJ463212 | ACCESSION:CJ463212 |
| c | 53 | 644.3 | 8.6 | 833 | 1 | BZ607553 | ACCESSION:BZ607553 |
| c | 54 | 642.6 | 8.6 | 744 | 1 | BZ601435 | ACCESSION:BZ601435 |
| | 55 | 629.6 | 8.4 | 701 | 1 | DA834293 | ACCESSION:DA834293 |
| | 56 | 629.6 | 8.4 | 791 | 1 | DA834387 | ACCESSION:DA834387 |
| | 57 | 627.1 | 8.4 | 851 | 1 | CR765248 | ACCESSION:CR765248 |
| | 58 | 626.7 | 8.4 | 724 | 1 | DA834354 | ACCESSION:DA834354 |
| c | 59 | 625.6 | 8.4 | 757 | 1 | CZ449910 | ACCESSION:CZ449910 |
| c | 60 | 622.1 | 8.3 | 658 | 1 | BX430050 | ACCESSION:BX430050 |
| | 61 | 620.7 | 8.3 | 733 | 1 | BX374329 | ACCESSION:BX374329 |
| | 62 | 618.2 | 8.3 | 738 | 1 | HSA408776 | ACCESSION:AJ408776 |
| c | 63 | 618 | 8.3 | 776 | 1 | AQ351025 | ACCESSION:AQ351025 |
| | 64 | 617.3 | 8.2 | 696 | 1 | DA830356 | ACCESSION:DA830356 |
| | 65 | 614.501 | 8.2 | 792 | 1 | BX409304 | ACCESSION:BX409304 |
| | 66 | 610.1 | 8.1 | 725 | 1 | BX352667 | ACCESSION:BX352667 |
| c | 67 | 601.8 | 8.0 | 750 | 1 | BZ597795 | ACCESSION:BZ597795 |
| | 68 | 596.101 | 8.0 | 790 | 1 | BG572445 | ACCESSION:BG572445 |
| | 69 | 594.9 | 7.9 | 723 | 1 | BI772113 | ACCESSION:BI772113 |
| c | 70 | 590.1 | 7.9 | 814 | 1 | BX430055 | ACCESSION:BX430055 |
| | 71 | 579.6 | 7.7 | 586 | 1 | BE019603 | ACCESSION:BE019603 |
| c | 72 | 577.8 | 7.7 | 677 | 1 | CZ459267 | ACCESSION:CZ459267 |
| | 73 | 576.2 | 7.7 | 631 | 1 | AQ261133 | ACCESSION:AQ261133 |
| | 74 | 573.5 | 7.7 | 690 | 1 | AU134310 | ACCESSION:AU134310 |
| | 75 | 569.4 | 7.6 | 571 | 1 | DA846799 | ACCESSION:DA846799 |
| c | 76 | 566.1 | 7.6 | 674 | 1 | AQ111469 | ACCESSION:AQ111469 |
| | 77 | 566 | 7.6 | 572 | 1 | DA851356 | ACCESSION:DA851356 |
| | 78 | 563.8 | 7.5 | 571 | 1 | DA846543 | ACCESSION:DA846543 |
| | 79 | 560.3 | 7.5 | 745 | 1 | AU121943 | ACCESSION:AU121943 |
| | 80 | 557.8 | 7.4 | 680 | 1 | BZ610514 | ACCESSION:BZ610514 |
| | 81 | 554.5 | 7.4 | 572 | 1 | DA851363 | ACCESSION:DA851363 |
| | 82 | 551.3 | 7.4 | 564 | 1 | DA836641 | ACCESSION:DA836641 |
| | 83 | 550.6 | 7.4 | 578 | 1 | DA845799 | ACCESSION:DA845799 |
| c | 84 | 550.5 | 7.4 | 665 | 1 | AG154259 | ACCESSION:AG154259 |
| | 85 | 548.4 | 7.3 | 665 | 1 | AQ021928 | ACCESSION:AQ021928 |
| | 86 | 548.3 | 7.3 | 567 | 1 | DA861649 | ACCESSION:DA861649 |
| | 87 | 547.5 | 7.3 | 562 | 1 | DA849456 | ACCESSION:DA849456 |
| | 88 | 543.8 | 7.3 | 547 | 1 | DA845887 | ACCESSION:DA845887 |
| c | 89 | 542.1 | 7.2 | 647 | 1 | AQ546582 | ACCESSION:AQ546582 |
| | 90 | 539.9 | 7.2 | 570 | 1 | DA847869 | ACCESSION:DA847869 |
| c | 91 | 539.8 | 7.2 | 587 | 1 | DA239644 | ACCESSION:DA239644 |
| | 92 | 537.6 | 7.2 | 540 | 1 | DA849357 | ACCESSION:DA849357 |
| | 93 | 537.4 | 7.2 | 541 | 1 | DA835080 | ACCESSION:DA835080 |
| | 94 | 536.8 | 7.2 | 647 | 1 | DU641525 | ACCESSION:DU641525 |
| | 95 | 535.7 | 7.2 | 563 | 1 | DA834637 | ACCESSION:DA834637 |
| | 96 | 535.2 | 7.1 | 538 | 1 | DA838045 | ACCESSION:DA838045 |
| | 97 | 534.6 | 7.1 | 543 | 1 | DB062480 | ACCESSION:DB062480 |
| | 98 | 534.2 | 7.1 | 539 | 1 | DB068120 | ACCESSION:DB068120 |
| | 99 | 533.1 | 7.1 | 711 | 1 | BX357209 | ACCESSION:BX357209 |
| | 100 | 532.5 | 7.1 | 547 | 1 | DA838367 | ACCESSION:DA838367 |
| | 101 | 531.5 | 7.1 | 568 | 1 | DA854810 | ACCESSION:DA854810 |
| | 102 | 531 | 7.1 | 569 | 1 | DA838561 | ACCESSION:DA838561 |
| c | 103 | 530.3 | 7.1 | 652 | 1 | BX952683 | ACCESSION:BX952683 |
| | 104 | 530 | 7.1 | 559 | 1 | DA856636 | ACCESSION:DA856636 |
| | 105 | 529.7 | 7.1 | 557 | 1 | DA858973 | ACCESSION:DA858973 |
| c | 106 | 524.9 | 7.0 | 541 | 1 | DB381575 | ACCESSION:DB381575 |

| | | | | | | |
|---|---|---|---|---|---|---|
| c 107 | 523.9 | 7.0 | 559 | 1 | AQ343877 | ACCESSION:AQ343877 |
| 108 | 520.5 | 6.9 | 552 | 1 | DA838102 | ACCESSION:DA838102 |
| 109 | 517 | 6.9 | 563 | 1 | DA375140 | ACCESSION:DA375140 |
| 110 | 515 | 6.9 | 554 | 1 | CR735592 | ACCESSION:CR735592 |
| c 111 | 507.7 | 6.8 | 594 | 1 | AQ193544 | ACCESSION:AQ193544 |
| 112 | 501.7 | 6.7 | 534 | 1 | DB028309 | ACCESSION:DB028309 |
| 113 | 501 | 6.7 | 503 | 1 | BG000981 | ACCESSION:BG000981 |
| c 114 | 500.3 | 6.7 | 531 | 1 | BG012022 | ACCESSION:BG012022 |
| 115 | 496.9 | 6.6 | 578 | 1 | DA848644 | ACCESSION:DA848644 |
| 116 | 491.7 | 6.6 | 558 | 1 | AQ012843 | ACCESSION:AQ012843 |
| 117 | 488.5 | 6.5 | 528 | 1 | BX492789 | ACCESSION:BX492789 |
| 118 | 488 | 6.5 | 610 | 1 | BF996761 | ACCESSION:BF996761 |
| 119 | 484 | 6.5 | 545 | 1 | DB066140 | ACCESSION:DB066140 |
| c 120 | 482.3 | 6.4 | 537 | 1 | DB381433 | ACCESSION:DB381433 |
| c 121 | 480.4 | 6.4 | 494 | 1 | AA781423 | ACCESSION:AA781423 |
| c 122 | 480.4 | 6.4 | 578 | 1 | AA613238 | ACCESSION:AA613238 |
| c 123 | 480.2 | 6.4 | 595 | 1 | B59017 | ACCESSION:B59017 |
| 124 | 480 | 6.4 | 631 | 1 | BG573364 | ACCESSION:BG573364 |
| c 125 | 478.9 | 6.4 | 544 | 1 | DB381156 | ACCESSION:DB381156 |
| 126 | 474.5 | 6.3 | 544 | 1 | DB068316 | ACCESSION:DB068316 |
| 127 | 474.3 | 6.3 | 569 | 1 | DB096981 | ACCESSION:DB096981 |
| c 128 | 472.1 | 6.3 | 526 | 1 | DB381477 | ACCESSION:DB381477 |
| c 129 | 471.4 | 6.3 | 494 | 1 | BQ365207 | ACCESSION:BQ365207 |
| c 130 | 470.1 | 6.3 | 598 | 1 | BZ611361 | ACCESSION:BZ611361 |
| 131 | 466.3 | 6.2 | 592 | 1 | DA673983 | ACCESSION:DA673983 |
| c 132 | 465.8 | 6.2 | 508 | 1 | DB381241 | ACCESSION:DB381241 |
| c 133 | 464.2 | 6.2 | 478 | 1 | DB378147 | ACCESSION:DB378147 |
| c 134 | 463.5 | 6.2 | 577 | 1 | DB377997 | ACCESSION:DB377997 |
| 135 | 462.5 | 6.2 | 491 | 1 | DA864742 | ACCESSION:DA864742 |
| c 136 | 461.8 | 6.2 | 521 | 1 | AZ517418 | ACCESSION:AZ517418 |
| c 137 | 456.3 | 6.1 | 533 | 1 | DB380923 | ACCESSION:DB380923 |
| c 138 | 454.2 | 6.1 | 507 | 1 | DB381613 | ACCESSION:DB381613 |
| c 139 | 449.8 | 6.0 | 564 | 1 | BX388917 | ACCESSION:BX388917 |
| c 140 | 448.7 | 6.0 | 625 | 1 | AQ541847 | ACCESSION:AQ541847 |
| c 141 | 444.9 | 5.9 | 520 | 1 | AG195724 | ACCESSION:AG195724 |
| 142 | 443.5 | 5.9 | 562 | 1 | DA835505 | ACCESSION:DA835505 |
| 143 | 439.7 | 5.9 | 572 | 1 | BQ365363 | ACCESSION:BQ365363 |
| c 144 | 439.3 | 5.9 | 509 | 1 | DB381029 | ACCESSION:DB381029 |
| 145 | 437.6 | 5.8 | 506 | 1 | AQ878055 | ACCESSION:AQ878055 |
| 146 | 436.7 | 5.8 | 565 | 1 | AQ553777 | ACCESSION:AQ553777 |
| 147 | 436.4 | 5.8 | 440 | 1 | BE732673 | ACCESSION:BE732673 |
| c 148 | 433.2 | 5.8 | 484 | 1 | BX409035 | ACCESSION:BX409035 |
| 149 | 433.2 | 5.8 | 596 | 1 | AQ384981 | ACCESSION:AQ384981 |
| 150 | 432.4 | 5.8 | 470 | 1 | BX492778 | ACCESSION:BX492778 |
| 151 | 431.2 | 5.8 | 560 | 1 | CZ465902 | ACCESSION:CZ465902 |
| 152 | 429.2 | 5.7 | 579 | 1 | DA045520 | ACCESSION:DA045520 |
| 153 | 427.8 | 5.7 | 490 | 1 | AA426511 | ACCESSION:AA426511 |
| c 154 | 425 | 5.7 | 427 | 1 | AA776439 | ACCESSION:AA776439 |
| c 155 | 422.7 | 5.6 | 498 | 1 | DB380906 | ACCESSION:DB380906 |
| c 156 | 421.3 | 5.6 | 450 | 1 | BQ024410 | ACCESSION:BQ024410 |
| c 157 | 420.6 | 5.6 | 533 | 1 | B17809 | ACCESSION:B17809 |
| c 158 | 420 | 5.6 | 460 | 1 | DN914892 | ACCESSION:DN914892 |
| 159 | 420 | 5.6 | 460 | 1 | DN914893 | ACCESSION:DN914893 |
| c 160 | 418.6 | 5.6 | 522 | 1 | AW971553 | ACCESSION:AW971553 |
| c 161 | 414.1 | 5.5 | 516 | 1 | DB322339 | ACCESSION:DB322339 |
| c 162 | 413.9 | 5.5 | 514 | 1 | DB377653 | ACCESSION:DB377653 |
| 163 | 413.8 | 5.5 | 539 | 1 | AQ042339 | ACCESSION:AQ042339 |
| c 164 | 413.6 | 5.5 | 568 | 1 | AQ568446 | ACCESSION:AQ568446 |
| c 165 | 412.8 | 5.5 | 490 | 1 | AI598135 | ACCESSION:AI598135 |
| c 166 | 409.5 | 5.5 | 439 | 1 | DB375490 | ACCESSION:DB375490 |
| c 167 | 408.9 | 5.5 | 537 | 1 | AQ270891 | ACCESSION:AQ270891 |
| c 168 | 407.2 | 5.4 | 414 | 1 | BF993783 | ACCESSION:BF993783 |
| c 169 | 406.8 | 5.4 | 436 | 1 | AI128526 | ACCESSION:AI128526 |
| 170 | 404.5 | 5.4 | 462 | 1 | BX953370 | ACCESSION:BX953370 |
| 171 | 402.9 | 5.4 | 479 | 1 | AQ226613 | ACCESSION:AQ226613 |
| 172 | 402.1 | 5.4 | 539 | 1 | AQ486506 | ACCESSION:AQ486506 |
| 173 | 401.4 | 5.4 | 423 | 1 | BF992535 | ACCESSION:BF992535 |
| c 174 | 399.9 | 5.3 | 557 | 1 | AQ061550 | ACCESSION:AQ061550 |
| 175 | 399.2 | 5.3 | 529 | 1 | AW967582 | ACCESSION:AW967582 |
| c 176 | 397.3 | 5.3 | 415 | 1 | AI128496 | ACCESSION:AI128496 |
| 177 | 395.2 | 5.3 | 534 | 1 | AQ483208 | ACCESSION:AQ483208 |
| c 178 | 395.1 | 5.3 | 543 | 1 | BI963185 | ACCESSION:BI963185 |
| c 179 | 394.1 | 5.3 | 527 | 1 | DB311329 | ACCESSION:DB311329 |
| 180 | 393.3 | 5.3 | 516 | 1 | AQ141344 | ACCESSION:AQ141344 |
| c 181 | 392.9 | 5.2 | 422 | 1 | N53177 | ACCESSION:N53177 |
| 182 | 392.9 | 5.2 | 501 | 1 | AG193108 | ACCESSION:AG193108 |
| c 183 | 392 | 5.2 | 549 | 1 | AQ231407 | ACCESSION:AQ231407 |
| c 184 | 391.9 | 5.2 | 458 | 1 | R76086 | ACCESSION:R76086 |
| c 185 | 391.9 | 5.2 | 542 | 1 | DB297732 | ACCESSION:DB297732 |
| c 186 | 385.2 | 5.1 | 542 | 1 | DB375050 | ACCESSION:DB375050 |
| 187 | 384.4 | 5.1 | 459 | 1 | R68740 | ACCESSION:R68740 |
| c 188 | 384 | 5.1 | 411 | 1 | BG004254 | ACCESSION:BG004254 |
| c 189 | 382.7 | 5.1 | 485 | 1 | AW511366 | ACCESSION:AW511366 |
| 190 | 379.8 | 5.1 | 433 | 1 | N77302 | ACCESSION:N77302 |
| 191 | 378.9 | 5.1 | 441 | 1 | DQ033443 | ACCESSION:DQ033443 |
| c 192 | 375.6 | 5.0 | 382 | 1 | BI052569 | ACCESSION:BI052569 |
| 193 | 375.5 | 5.0 | 466 | 1 | B38783 | ACCESSION:B38783 |
| c 194 | 374.4 | 5.0 | 431 | 1 | AQ221075 | ACCESSION:AQ221075 |
| c 195 | 374.2 | 5.0 | 376 | 1 | BI056595 | ACCESSION:BI056595 |
| 196 | 373.3 | 5.0 | 521 | 1 | AQ536209 | ACCESSION:AQ536209 |
| 197 | 372.7 | 5.0 | 434 | 1 | AQ167811 | ACCESSION:AQ167811 |
| 198 | 372.4 | 5.0 | 379 | 1 | BF993049 | ACCESSION:BF993049 |
| c 199 | 372.2 | 5.0 | 373 | 1 | BI052559 | ACCESSION:BI052559 |
| c 200 | 372.2 | 5.0 | 373 | 1 | BI052567 | ACCESSION:BI052567 |
| c 201 | 371.6 | 5.0 | 374 | 1 | BI056593 | ACCESSION:BI056593 |
| c 202 | 370.7 | 4.9 | 449 | 1 | AQ734945 | ACCESSION:AQ734945 |
| c 203 | 369 | 4.9 | 373 | 1 | BI052561 | ACCESSION:BI052561 |
| c 204 | 369 | 4.9 | 373 | 1 | BI052571 | ACCESSION:BI052571 |
| c 205 | 369 | 4.9 | 373 | 1 | BI052572 | ACCESSION:BI052572 |
| c 206 | 369 | 4.9 | 373 | 1 | BI056589 | ACCESSION:BI056589 |
| 207 | 368.5 | 4.9 | 467 | 1 | AQ064818 | ACCESSION:AQ064818 |
| 208 | 368.1 | 4.9 | 443 | 1 | AQ057826 | ACCESSION:AQ057826 |
| 209 | 367.9 | 4.9 | 423 | 1 | AQ034878 | ACCESSION:AQ034878 |
| c 210 | 367.4 | 4.9 | 373 | 1 | BI052570 | ACCESSION:BI052570 |
| c 211 | 367.2 | 4.9 | 379 | 1 | BI052566 | ACCESSION:BI052566 |
| 212 | 366.1 | 4.9 | 408 | 1 | AQ216381 | ACCESSION:AQ216381 |
| c 213 | 365.5 | 4.9 | 443 | 1 | AA837267 | ACCESSION:AA837267 |
| c 214 | 363.2 | 4.8 | 388 | 1 | H01325 | ACCESSION:H01325 |
| c 215 | 363.2 | 4.8 | 470 | 1 | AI074704 | ACCESSION:AI074704 |
| c 216 | 360.8 | 4.8 | 438 | 1 | R77278 | ACCESSION:R77278 |
| c 217 | 359.5 | 4.8 | 446 | 1 | CB069106 | ACCESSION:CB069106 |
| c 218 | 358.7 | 4.8 | 399 | 1 | DB374307 | ACCESSION:DB374307 |
| c 219 | 358 | 4.8 | 442 | 1 | AQ701392 | ACCESSION:AQ701392 |
| c 220 | 356.2 | 4.8 | 446 | 1 | AI393478 | ACCESSION:AI393478 |
| 221 | 356.1 | 4.8 | 476 | 1 | AQ774432 | ACCESSION:AQ774432 |
| 222 | 355.8 | 4.8 | 372 | 1 | BF992537 | ACCESSION:BF992537 |
| c 223 | 355.6 | 4.7 | 425 | 1 | AI570707 | ACCESSION:AI570707 |
| c 224 | 355.6 | 4.7 | 491 | 1 | B17135 | ACCESSION:B17135 |
| c 225 | 355.3 | 4.7 | 431 | 1 | AA552941 | ACCESSION:AA552941 |
| 226 | 351.4 | 4.7 | 412 | 1 | AA251070 | ACCESSION:AA251070 |
| 227 | 350.7 | 4.7 | 399 | 1 | DQ033444 | ACCESSION:DQ033444 |
| c 228 | 350.7 | 4.7 | 456 | 1 | AQ232635 | ACCESSION:AQ232635 |
| 229 | 350.1 | 4.7 | 440 | 1 | AG196511 | ACCESSION:AG196511 |
| 230 | 350 | 4.7 | 366 | 1 | BF993782 | ACCESSION:BF993782 |
| 231 | 349.3 | 4.7 | 472 | 1 | H01372 | ACCESSION:H01372 |
| c 232 | 348.7 | 4.7 | 481 | 1 | AQ609759 | ACCESSION:AQ609759 |
| c 233 | 348.6 | 4.7 | 428 | 1 | AQ132081 | ACCESSION:AQ132081 |
| 234 | 348.6 | 4.7 | 464 | 1 | R78828 | ACCESSION:R78828 |
| 235 | 345.5 | 4.6 | 432 | 1 | AV708482 | ACCESSION:AV708482 |
| 236 | 345 | 4.6 | 379 | 1 | D78692 | ACCESSION:D78692 |
| c 237 | 344.4 | 4.6 | 433 | 1 | AI379210 | ACCESSION:AI379210 |
| 238 | 341.8 | 4.6 | 468 | 1 | N41878 | ACCESSION:N41878 |
| c 239 | 341.6 | 4.6 | 424 | 1 | R27412 | ACCESSION:R27412 |
| 240 | 340.8 | 4.6 | 408 | 1 | AQ210561 | ACCESSION:AQ210561 |
| 241 | 339.7 | 4.5 | 402 | 1 | AQ734369 | ACCESSION:AQ734369 |
| 242 | 339.5 | 4.5 | 362 | 1 | BI049622 | ACCESSION:BI049622 |
| 243 | 339.5 | 4.5 | 378 | 1 | R76136 | ACCESSION:R76136 |
| c 244 | 339 | 4.5 | 385 | 1 | N55091 | ACCESSION:N55091 |
| c 245 | 339 | 4.5 | 410 | 1 | AA250958 | ACCESSION:AA250958 |
| c 246 | 338.8 | 4.5 | 403 | 1 | BX388918 | ACCESSION:BX388918 |
| c 247 | 338.4 | 4.5 | 345 | 1 | BI056592 | ACCESSION:BI056592 |
| c 248 | 338.1 | 4.5 | 446 | 1 | AI288235 | ACCESSION:AI288235 |
| c 249 | 337.3 | 4.5 | 442 | 1 | AQ075734 | ACCESSION:AQ075734 |
| c 250 | 337.3 | 4.5 | 467 | 1 | AQ876467 | ACCESSION:AQ876467 |
| c 251 | 336.6 | 4.5 | 431 | 1 | AQ736815 | ACCESSION:AQ736815 |
| c 252 | 336.2 | 4.5 | 463 | 1 | R68685 | ACCESSION:R68685 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 253 | 332.6 | 4.4 | 408 | 1 | BU790503 | ACCESSION:BU790503 |
| 254 | 332.6 | 4.4 | 408 | 1 | CA848433 | ACCESSION:CA848433 |
| c 255 | 332.3 | 4.4 | 429 | 1 | AQ028930 | ACCESSION:AQ028930 |
| 256 | 332.2 | 4.4 | 367 | 1 | BF990707 | ACCESSION:BF990707 |
| 257 | 331 | 4.4 | 408 | 1 | BU790429 | ACCESSION:BU790429 |
| 258 | 330.7 | 4.4 | 397 | 1 | AQ242769 | ACCESSION:AQ242769 |
| c 259 | 329.4 | 4.4 | 346 | 1 | BI053386 | ACCESSION:BI053386 |
| 260 | 329.2 | 4.4 | 430 | 1 | AQ010439 | ACCESSION:AQ010439 |
| c 261 | 329.1 | 4.4 | 357 | 1 | DB375009 | ACCESSION:DB375009 |
| c 262 | 326.7 | 4.4 | 461 | 1 | AQ225099 | ACCESSION:AQ225099 |
| c 263 | 326.4 | 4.4 | 363 | 1 | T69704 | ACCESSION:T69704 |
| c 264 | 324.7 | 4.3 | 438 | 1 | CA432887 | ACCESSION:CA432887 |
| 265 | 323.9 | 4.3 | 354 | 1 | BQ365139 | ACCESSION:BQ365139 |
| 266 | 323.7 | 4.3 | 418 | 1 | B95002 | ACCESSION:B95002 |
| c 267 | 323 | 4.3 | 342 | 1 | AA860368 | ACCESSION:AA860368 |
| c 268 | 321.6 | 4.3 | 346 | 1 | BQ365143 | ACCESSION:BQ365143 |
| 269 | 320.9 | 4.3 | 430 | 1 | AQ345621 | ACCESSION:AQ345621 |
| c 270 | 320.8 | 4.3 | 342 | 1 | T47345 | ACCESSION:T47345 |
| 271 | 320.6 | 4.3 | 416 | 1 | AQ200416 | ACCESSION:AQ200416 |
| c 272 | 320.6 | 4.3 | 449 | 1 | AQ701255 | ACCESSION:AQ701255 |
| 273 | 317.3 | 4.2 | 422 | 1 | AQ150723 | ACCESSION:AQ150723 |
| c 274 | 317 | 4.2 | 404 | 1 | BU790130 | ACCESSION:BU790130 |
| c 275 | 315.5 | 4.2 | 420 | 1 | DB352776 | ACCESSION:DB352776 |
| 276 | 314.7 | 4.2 | 421 | 1 | BF761543 | ACCESSION:BF761543 |
| 277 | 313.6 | 4.2 | 399 | 1 | AQ005223 | ACCESSION:AQ005223 |
| c 278 | 312.1 | 4.2 | 385 | 1 | AQ062249 | ACCESSION:AQ062249 |
| 279 | 312 | 4.2 | 377 | 1 | AQ044774 | ACCESSION:AQ044774 |
| 280 | 311.9 | 4.2 | 413 | 1 | AQ040947 | ACCESSION:AQ040947 |
| c 281 | 311.9 | 4.2 | 427 | 1 | AA577847 | ACCESSION:AA577847 |
| 282 | 311.4 | 4.2 | 399 | 1 | AG195659 | ACCESSION:AG195659 |
| c 283 | 310.1 | 4.1 | 334 | 1 | BI053385 | ACCESSION:BI053385 |
| 284 | 309 | 4.1 | 316 | 1 | BI053220 | ACCESSION:BI053220 |
| c 285 | 308.9 | 4.1 | 398 | 1 | AQ044697 | ACCESSION:AQ044697 |
| c 286 | 308.8 | 4.1 | 437 | 1 | B40768 | ACCESSION:B40768 |
| c 287 | 308.1 | 4.1 | 408 | 1 | AL598182 | ACCESSION:AL598182 |
| 288 | 304.1 | 4.1 | 364 | 1 | AA368311 | ACCESSION:AA368311 |
| c 289 | 303.7 | 4.1 | 366 | 1 | AQ549534 | ACCESSION:AQ549534 |
| c 290 | 301.6 | 4.0 | 313 | 1 | BI056596 | ACCESSION:BI056596 |
| 291 | 300.8 | 4.0 | 342 | 1 | AA663366 | ACCESSION:AA663366 |
| c 292 | 300.8 | 4.0 | 404 | 1 | R27389 | ACCESSION:R27389 |
| c 293 | 300.5 | 4.0 | 383 | 1 | AQ192721 | ACCESSION:AQ192721 |
| 294 | 300.5 | 4.0 | 417 | 1 | AA781468 | ACCESSION:AA781468 |
| 295 | 300.2 | 4.0 | 414 | 1 | H12439 | ACCESSION:H12439 |
| 296 | 299.9 | 4.0 | 398 | 1 | AQ412083 | ACCESSION:AQ412083 |
| 297 | 298.8 | 4.0 | 322 | 1 | BF990902 | ACCESSION:BF990902 |
| c 298 | 297.6 | 4.0 | 310 | 1 | BQ365206 | ACCESSION:BQ365206 |
| c 299 | 297.2 | 4.0 | 298 | 1 | BF992540 | ACCESSION:BF992540 |
| 300 | 295.7 | 3.9 | 407 | 1 | R27388 | ACCESSION:R27388 |
| 301 | 294.7 | 3.9 | 358 | 1 | AQ020123 | ACCESSION:AQ020123 |
| c 302 | 293.6 | 3.9 | 328 | 1 | AA729556 | ACCESSION:AA729556 |
| 303 | 293.5 | 3.9 | 377 | 1 | AQ033092 | ACCESSION:AQ033092 |
| 304 | 290.8 | 3.9 | 374 | 1 | T93661 | ACCESSION:T93661 |
| c 305 | 289 | 3.9 | 332 | 1 | AQ024733 | ACCESSION:AQ024733 |
| c 306 | 288.1 | 3.8 | 347 | 1 | BQ024998 | ACCESSION:BQ024998 |
| c 307 | 286.4 | 3.8 | 315 | 1 | BI055459 | ACCESSION:BI055459 |
| c 308 | 286.4 | 3.8 | 315 | 1 | BI056080 | ACCESSION:BI056080 |
| c 309 | 285.9 | 3.8 | 331 | 1 | BI055458 | ACCESSION:BI055458 |
| c 310 | 285.9 | 3.8 | 331 | 1 | BI056079 | ACCESSION:BI056079 |
| c 311 | 284.8 | 3.8 | 300 | 1 | BI055456 | ACCESSION:BI055456 |
| c 312 | 284.8 | 3.8 | 300 | 1 | BI056077 | ACCESSION:BI056077 |
| 313 | 283.6 | 3.8 | 289 | 1 | BI053219 | ACCESSION:BI053219 |
| 314 | 283.3 | 3.8 | 381 | 1 | B63737 | ACCESSION:B63737 |
| 315 | 278.6 | 3.7 | 302 | 1 | BG004247 | ACCESSION:BG004247 |
| 316 | 276.6 | 3.7 | 326 | 1 | D29167 | ACCESSION:D29167 |
| 317 | 276.4 | 3.7 | 372 | 1 | T70359 | ACCESSION:T70359 |
| c 318 | 274.7 | 3.7 | 364 | 1 | B59529 | ACCESSION:B59529 |
| 319 | 274.4 | 3.7 | 317 | 1 | BI056590 | ACCESSION:BI056590 |
| 320 | 270.5 | 3.6 | 353 | 1 | BM721661 | ACCESSION:BM721661 |
| c 321 | 270 | 3.6 | 320 | 1 | BX388919 | ACCESSION:BX388919 |
| c 322 | 267.7 | 3.6 | 373 | 1 | R31587 | ACCESSION:R31587 |
| c 323 | 267.1 | 3.6 | 345 | 1 | AA868777 | ACCESSION:AA868777 |
| c 324 | 264.8 | 3.5 | 370 | 1 | AQ001683 | ACCESSION:AQ001683 |
| c 325 | 260.4 | 3.5 | 328 | 1 | AQ079098 | ACCESSION:AQ079098 |
| 326 | 260 | 3.5 | 352 | 1 | AA179079 | ACCESSION:AA179079 |
| c 327 | 258.8 | 3.5 | 331 | 1 | BF943201 | ACCESSION:BF943201 |
| c 328 | 258.7 | 3.5 | 330 | 1 | AA774109 | ACCESSION:AA774109 |
| 329 | 258.3 | 3.4 | 333 | 1 | AQ030065 | ACCESSION:AQ030065 |
| 330 | 257.5 | 3.4 | 344 | 1 | T52094 | ACCESSION:T52094 |
| c 331 | 256.5 | 3.4 | 327 | 1 | AG266374 | ACCESSION:AG266374 |
| c 332 | 254.3 | 3.4 | 332 | 1 | AI797629 | ACCESSION:AI797629 |
| c 333 | 254.3 | 3.4 | 351 | 1 | AQ058742 | ACCESSION:AQ058742 |
| 334 | 252.4 | 3.4 | 349 | 1 | AQ116788 | ACCESSION:AQ116788 |
| c 335 | 252 | 3.4 | 293 | 1 | H04137 | ACCESSION:H04137 |
| c 336 | 251.9 | 3.4 | 327 | 1 | AQ241082 | ACCESSION:AQ241082 |
| c 337 | 250.2 | 3.3 | 311 | 1 | AQ007989 | ACCESSION:AQ007989 |
| 338 | 247.4 | 3.3 | 299 | 1 | BI007745 | ACCESSION:BI007745 |
| 339 | 245.8 | 3.3 | 305 | 1 | BF990710 | ACCESSION:BF990710 |
| c 340 | 244.7 | 3.3 | 335 | 1 | R27591 | ACCESSION:R27591 |
| c 341 | 243.5 | 3.3 | 334 | 1 | BQ353502 | ACCESSION:BQ353502 |
| 342 | 242.4 | 3.2 | 283 | 1 | T87403 | ACCESSION:T87403 |
| c 343 | 242.2 | 3.2 | 292 | 1 | CA848155 | ACCESSION:CA848155 |
| 344 | 240.1 | 3.2 | 259 | 1 | R37932 | ACCESSION:R37932 |
| c 345 | 239.4 | 3.2 | 272 | 1 | T87297 | ACCESSION:T87297 |
| c 346 | 239 | 3.2 | 323 | 1 | BG014730 | ACCESSION:BG014730 |
| c 347 | 237.3 | 3.2 | 323 | 1 | R00744 | ACCESSION:R00744 |
| c 348 | 232.3 | 3.1 | 322 | 1 | BG015482 | ACCESSION:BG015482 |
| c 349 | 230.7 | 3.1 | 310 | 1 | AQ000990 | ACCESSION:AQ000990 |
| c 350 | 230.4 | 3.1 | 291 | 1 | BF989696 | ACCESSION:BF989696 |
| 351 | 230.1 | 3.1 | 269 | 1 | DU131223 | ACCESSION:DU131223 |
| c 352 | 230 | 3.1 | 314 | 1 | AQ539232 | ACCESSION:AQ539232 |
| c 353 | 229.6 | 3.1 | 278 | 1 | BU790206 | ACCESSION:BU790206 |
| 354 | 223.8 | 3.0 | 244 | 1 | BQ365242 | ACCESSION:BQ365242 |
| c 355 | 221.9 | 3.0 | 268 | 1 | B46293 | ACCESSION:B46293 |
| 356 | 221.6 | 3.0 | 293 | 1 | BE082150 | ACCESSION:BE082150 |
| 357 | 220.9 | 2.9 | 242 | 1 | R31328 | ACCESSION:R31328 |
| c 358 | 220.8 | 2.9 | 290 | 1 | BQ022458 | ACCESSION:BQ022458 |
| 359 | 219.8 | 2.9 | 267 | 1 | H13097 | ACCESSION:H13097 |
| c 360 | 214.4 | 2.9 | 239 | 1 | AQ264399 | ACCESSION:AQ264399 |
| 361 | 214.1 | 2.9 | 238 | 1 | BQ365241 | ACCESSION:BQ365241 |
| c 362 | 213.9 | 2.9 | 239 | 1 | H13302 | ACCESSION:H13302 |
| 363 | 209.3 | 2.8 | 294 | 1 | R01948 | ACCESSION:R01948 |
| c 364 | 206.8 | 2.8 | 289 | 1 | B33508 | ACCESSION:B33508 |
| 365 | 195.8 | 2.6 | 237 | 1 | BF954175 | ACCESSION:BF954175 |
| c 366 | 187.8 | 2.5 | 195 | 1 | BF992685 | ACCESSION:BF992685 |
| c 367 | 183.5 | 2.5 | 229 | 1 | AQ040405 | ACCESSION:AQ040405 |
| c 368 | 177.4 | 2.4 | 235 | 1 | BF810306 | ACCESSION:BF810306 |
| c 369 | 175.2 | 2.3 | 198 | 1 | DU639856 | ACCESSION:DU639856 |
| c 370 | 175.2 | 2.3 | 198 | 1 | DU640205 | ACCESSION:DU640205 |
| 371 | 165.6 | 2.2 | 198 | 1 | AA077794 | ACCESSION:AA077794 |
| c 372 | 156.5 | 2.1 | 213 | 1 | AQ027014 | ACCESSION:AQ027014 |
| 373 | 153.5 | 2.0 | 190 | 1 | DQ032120 | ACCESSION:DQ032120 |
| 374 | 150.3 | 2.0 | 190 | 1 | DQ032119 | ACCESSION:DQ032119 |
| c 375 | 145.9 | 1.9 | 204 | 1 | AI611840 | ACCESSION:AI611840 |
| c 376 | 145.4 | 1.9 | 183 | 1 | AQ479009 | ACCESSION:AQ479009 |
| 377 | 143.7 | 1.9 | 171 | 1 | BG013341 | ACCESSION:BG013341 |
| 378 | 141.4 | 1.9 | 173 | 1 | B57639 | ACCESSION:B57639 |
| c 379 | 131.8 | 1.8 | 140 | 1 | CB298587 | ACCESSION:CB298587 |
| 380 | 127.6 | 1.7 | 128 | 1 | T39519 | ACCESSION:T39519 |
| c 381 | 123.6 | 1.7 | 146 | 1 | CR958717 | ACCESSION:CR958717 |
| 382 | 100.2 | 1.3 | 118 | 1 | AG197813 | ACCESSION:AG197813 |
| c 383 | 77 | 1.0 | 107 | 1 | BQ368374 | ACCESSION:BQ368374 |

ALIGNMENTS

```
RESULT 1
CR605851
LOCUS       CR605851               2748 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens
            (human).
ACCESSION   CR605851
VERSION     CR605851.1  GI:50486658
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2748)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 2748)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .2748
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE012YJ24"
                     /tissue_type="Placenta"
                     /plasmid="pCMVSPORT_6"

  Query Match             29.1%;  Score 2180.9;  DB 1;  Length 2748;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches 2171;  Conservative   34;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        544 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 603

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 663

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 723

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| |||||||:||||||
Db        724 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 782

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        783 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 842

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        843 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 902

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        903 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 962

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1022

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1082

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1142

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db       1143 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1202

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1262

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1322

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       1323 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1382

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1442

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1502

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1562

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1563 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1622

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||
Db       1623 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1682

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1742

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1743 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1802

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1862

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1863 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1922

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1923 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1982

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1983 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2042

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:|||||:|||||||||||||||||||||||||||||||||||||:|||
Db       2043 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2102

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2103 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2162

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||:||
Db       2163 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2222
```

```
Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2282

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2283 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2342

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2343 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2402

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2403 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2462

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2463 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2522

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2523 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2582

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:|||||:||||||||||||
Db       2583 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2642

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2643 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2702

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTT 7557
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       2703 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTT 2748
```

RESULT 2
CR617248
LOCUS CR617248 2749 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI022YJ18 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617248
VERSION CR617248.1 GI:50498055
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2749)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI022YJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 29.1%; Score 2176.9; DB 1; Length 2749;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2167; Conservative 34; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        549 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 608

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 668

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 728

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        729 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 787

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db        788 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 847

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        848 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 907

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        908 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 967

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1027

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1087

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1147

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1148 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1207

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1267

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1327

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       1328 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1387

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1447

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1507

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
```

```
Db       1508 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1567

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1568 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1627

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db       1628 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1687

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1688 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1747

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1748 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1807

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1808 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1867

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1868 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1927

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1928 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1987

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||||
Db       1988 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2047

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||||:|||
Db       2048 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2107

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2167

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||||
Db       2168 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2227

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2287

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2347

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2348 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2407

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2408 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2467

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||
Db       2468 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2527

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2528 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2587

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||

Db       2588 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2647

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2648 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2707

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCT 7553
             |||||||||||||||||||||||||||||||||||||||||||
Db       2708 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCT 2749
```

RESULT 3
CR622175
LOCUS CR622175 2500 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR622175
VERSION CR622175.1 GI:50502982
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2500)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2500)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 29.0%; Score 2174.9; DB 1; Length 2500;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2165; Conservative 34; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        302 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 361

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 421

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 481

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| |||||||:|||||||
Db        482 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 540

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||
```

```
Db        541 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 600

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
              ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        601 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 660

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        661 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 720

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 780

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 840

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 900

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        901 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 960

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1020

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1080

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1081 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1140

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1200

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1260

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1320

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1380

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
              ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
Db       1381 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1440

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1500

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1560

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1620

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1621 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1680

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
              ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1681 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1740

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||:|||||
Db       1741 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 1800

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
              ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||
Db       1801 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 1860

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 1920

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||:||||||
Db       1921 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 1980

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2040

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2100

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2101 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2160

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2220

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
              ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||
Db       2221 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2280

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2281 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2340

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
              ||||||||||||||||||||||||||||||||||||||||||:|||||:||||||||||
Db       2341 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2400

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db       2401 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2460

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 7551
              |||||||||||||||||||||||||||||||||||||||
Db       2461 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 2500
```

```
RESULT 4
CR625046
LOCUS       CR625046               2748 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI044YK06 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR625046
VERSION     CR625046.1  GI:50505853
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2748)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 2748)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .2748
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI044YK06"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

  Query Match             29.0%;  Score 2169.9;  DB 1;  Length 2748;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches 2160;  Conservative   34;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        555 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 614

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 674

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 734

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        735 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 793

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        794 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 853

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        854 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 913

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        914 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 973

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1033

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1093

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1153

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1154 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1213

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1273

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1333

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1334 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1393

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1453

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1513

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1573

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1633

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||
Db       1634 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1693

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1753

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1813

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1814 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1873

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1874 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1933

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       1934 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1993

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1994 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2053

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db       2054 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2113

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2114 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2173

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||
Db       2174 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2233

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2234 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2293

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
```

```
                |||||:||||||
```

```
Db       1564 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1623

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
              ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1624 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1683

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1684 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1743

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1803

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1863

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1864 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1923

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
              ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||
Db       1924 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1983

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1984 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2043

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
              ||||||||||:||||||:||||||||||||||||||||||||||||||||||||||:|||
Db       2044 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2103

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2163

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||
Db       2164 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2223

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2224 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2283

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2284 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2343

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2344 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2403

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2404 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2463

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
              ||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2464 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2523

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2524 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2583

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
              ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||
Db       2584 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2643

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2644 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2703

Qy       7512 TTTGGGTCCCCTC 7524
              |||||||||||||
Db       2704 TTTGGGTCCCCTC 2716


RESULT 6
BX365066/c
LOCUS       BX365066               1071 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI046YA18 3-PRIME, mRNA sequence.
ACCESSION   BX365066
VERSION     BX365066.2  GI:46304105
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1071)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30374869.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS1AI012ZC10NP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .1071
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI046YA18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             13.0%;  Score 970;  DB 1;  Length 1071;
  Best Local Similarity   95.8%;  Pred. No. 0;
  Matches  985;  Conservative   25;  Mismatches   16;  Indels    2;  Gaps    2;

Qy       6520 CGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTA 6579
              | |  || || ||   ||  |: |||:|||:|||||||||||||||||| ||||||||
Db       1026 CCCCSCACAAAAGAGWCCCATYCTCCYTTTKTTATAGGAGCAGGAGTGCTA-GTGCACTA 968

Qy       6580 GGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAA 6639
              || |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        967 GGGACTGGCATTGGCGGTATAACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAA 908

Qy       6640 CTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAAC 6699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 CTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAAC 848

Qy       6700 TCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGA 6759
              ||||||||||||||||||||:||||||||||||||||||||:||||||||||||:|||
Db        847 TCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGA 788

Qy       6760 GGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATC 6819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 GGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATC 728
```

```
Qy       6820 GTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGA 6879
             ||||||||||||||:|||||||||:||||||||||||||||:||||||:|||||||||
Db        727 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGA 668

Qy       6880 AACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGA 6939
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        667 AACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGA 608

Qy       6940 CCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTT 6999
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        607 CCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTT 548

Qy       7000 AACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCC 7059
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        547 AACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCC 488

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        487 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAAT 428

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 368

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        367 TCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 308

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        307 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 248

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:||||||||||||||||||||||||||||||||||||||||||||||:|
Db        247 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCT 188

Qy       7360 CACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAG 7419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 CACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAG 128

Qy       7420 AAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAA 7479
             ||||||||||||||:|||||:||||||||||||||||||||||||||||||||||||||
Db        127 AAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAA 68

Qy       7480 CCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGC 7539
             |||||||:|||||||||||||||||||||||||||||  |||||||| |||||||  ||::
Db         67 CCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGTNCCCCTCCC-TTGTATGCGASM 9

Qy       7540 TCTGTTTT 7547
              ||:|||||
Db           8 TCKGTTTT 1
```

RESULT 7
BX337769/c

```
LOCUS       BX337769                998 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION   BX337769
VERSION     BX337769.2  GI:46272079
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30337641.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
```

Query Match 12.3%; Score 923.7; DB 1; Length 998;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 956; Conservative 26; Mismatches 11; Indels 5; Gaps 4;

```
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACC-TCTACTCAGTT 6617
             ||||||||:||||||:|||||||  |||:||||||  ||||||||||| ||: |||||||
Db         998 GCAGGAGTKCTAGGTSCACTAGGGACTKGCATTGCAGGTATCACAACCTTCWCCTCAGTT 939

Qy       6618 CTACTACAAACTATCTCAAGAACT-AAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         938 CTACTACAAACTATCTCAAGAACTAAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 879

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||:|||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db         878 TCACCTKGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 819

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db         818 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 759

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db         758 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 699

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||||
Db         698 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 639

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 579

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||
Db         578 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 519

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         518 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 459

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAA-GGCACCCCTCCTGAGGAAATCTCAGCT 7155
             :|:|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         458 CTRGCCCACGATCTGATGTTAATGACATCAAACGGCACCCCTCCTGAGGAAATCTCAGCT 399

Qy       7156 GCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACC 7215
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db         398 GCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACC 339

Qy       7216 TCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGG 7275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        338 TCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGG 279

Qy       7276 ATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTT 7335
             |||||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||
Db        278 ATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTT 219

Qy       7336 AAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTA 7395
             ||||||||||||||||||||:|:||||||||||||||||||||||||||||||||||||||
Db        218 AAACACGGGGCTTGCAACTTRGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTA 159

Qy       7396 ATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAG 7455
             ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db        158 ATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAG 99

Qy       7456 GAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTG 7515
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         98 GAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTG 39

Qy       7516 GGTCCC--CTCCCTTTGTATGGGAGCTCTGTTTTCATG 7551
             ||||||  | ||  ||    ||||||  ||||||||||||
Db         38 GGTCCCNNCNCCTTTGTANTGGGAGGTCTGTTTTCATG 1
```

RESULT 8
BX378303/c

```
LOCUS       BX378303                1058 bp    mRNA    linear   EST 23-APR-2004
DEFINITION  BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI018YH07 3-PRIME, mRNA sequence.
ACCESSION   BX378303
VERSION     BX378303.2  GI:46557492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1058)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30439129.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI018CD04NP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .1058
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI018YH07"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
  Query Match             12.2%;  Score 912.6;  DB 1;  Length 1058;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  948;  Conservative   37;  Mismatches   49;  Indels    6;  Gaps    3;

Qy       6508 ATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTG 6567
             || : :|    :| |   |:|| : |||||||||| | ||:|  ||   |||||||||:|
Db       1038 ATGYAWATCAASCCCSCAAMAAARATACCCATTCTCCTTTKTTATA---GAGCAGGAGKG 982

Qy       6568 CTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAA 6627
             ||||  :      |||||||||||||| ||||||||||||||||||||||||||||||||
Db        981 CTAG--KGCACTAGTACTGGCATTGGC-GTATCACAACCTCTACTCAGTTCTACTACAAA 925

Qy       6628 CTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAA 6687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAA 865

Qy       6688 GATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTA 6747
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||:||||
Db        864 GATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTA 805

Qy       6748 ACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAAT 6807
             ||||||||:|||||:::|:|:| ||| |||||||||||||||||||||| |||||||||
Db        804 ACCGCTGAAAGAGGKKKAWCYTTTTTTTTTTTTAGGGGAAGAATGCTGTTACTATGTTAAT 745

Qy       6808 CAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCA 6867
             ||||||::|:||:|||||||||||||||:|||||||||:|||||||||||||||||:||||
Db        744 CAATCCKKAWTCKTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCA 685

Qy       6868 GARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTC 6927
             ||:||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        684 GAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTC 625

Qy       6928 CCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTT 6987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 CCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTT 565

Qy       6988 RACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCC 7047
             :|||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        564 AACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCC 505

Qy       7048 AAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGA 7107
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        504 AAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGA 445

Qy       7108 TCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTA 7167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTA 385

Qy       7168 CTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCA 7227
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        384 CTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCA 325

Qy       7228 CTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCT 7287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        324 CTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCYTAGGCT 265

Qy       7288 GAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCT 7347
             ||:|||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||||
Db        264 GACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCT 205

Qy       7348 TGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGA 7407
             ||||||||||:|||||||||||||||||||||||||||||||||||||| ||||||||||
Db        204 TGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAACTAGGCAAAGA 145

Qy       7408 CAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGA 7467
             |||||||||||||||||||||||||||||:||||||:|||||||||||||||||||||||||
Db        144 CAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGA 85

Qy       7468 TCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCT 7527
             ||||||||||||||||||:|||||||||||||||||||||||  |  |||  |
Db         84 TCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCNNTNNTGGTNNCNNTGGC 25

Qy       7528 TTGTATGGGAGCTCTGTTTT 7547
              | | || |   | |  |||
Db         24 GTTTGTGTGGAGTTTTGTTT 5
```

RESULT 9
BX391741/c

```
LOCUS       BX391741                935 bp    mRNA    linear   EST 29-APR-2004
```

```
DEFINITION  BX391741 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION   BX391741
VERSION     BX391741.1  GI:30619529
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAI022ZA10_CS02089_1&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             12.1%;  Score 909.1;  DB 1;  Length 935;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  919;  Conservative    6;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       5805 TGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAG 5864
             |||| |||||||||||||||||||||||| |||||||||| |  ||||||||||||||||
Db        933 TGCCCCTCTTTGCATGCATGCAAATACTCA-TATTGGACAGAANNAATGATTAATCCTAG 875

Qy       5865 TTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTC 5924
             |||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        874 TTGTCCTGNAGGACTTGAAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTC 815

Qy       5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCA 5984
             |||||||||||||||||||||||||||||| |||||||||||||||||| |||||||
Db        814 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATGTAAAAGAAGTATTCTCCCA 755

Qy       5985 ACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 ACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACA 695

Qy       6045 TGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCT 6104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 TGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCT 635

Qy       6105 CCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAR 6164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        634 CCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAG 575

Qy       6165 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 6224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 515

Qy       6225 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCT 6284
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        514 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCT 455

Qy       6285 CACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 6344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 CACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 395

Qy       6345 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 6404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 335

Qy       6405 AGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCC 6464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 AGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCC 275

Qy       6465 CCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
             |||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 215

Qy       6525 CAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTAC 6584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 CAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTAC 155

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 95

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          94 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 35

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||:||||||||||||||||||
Db          34 AGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 1


RESULT 10
BX347952
LOCUS       BX347952                955 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX347952 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION   BX347952
VERSION     BX347952.1  GI:30375235
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAE023ZF07_AE02167_1&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
```

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 11.8%; Score 887.2; DB 1; Length 955;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 931; Conservative 8; Mismatches 13; Indels 4; Gaps 4;

```
Qy       5806 GCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGT 5865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGT 60

Qy       5866 TGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCT 5925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCT 120

Qy       5926 GATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAA 5985
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 GATGGGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATGTAAAAGAAGTAATCTCCCAA 180

Qy       5986 CTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACAT 6045
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACAT 240

Qy       6046 GAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTC 6105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTC 300

Qy       6106 CATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARG 6165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        301 CATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGG 360

Qy       6166 CCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 6225
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 420

Qy       6226 ACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTC 6285
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        421 ACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTC 480

Qy       6286 ACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTA 6345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTA 540

Qy       6346 ACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCA 6405
             |||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTCCT-CCAAACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCA 599

Qy       6406 GCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCC 6465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCC 659

Qy       6466 CCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAAC 6525
             ||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAAC 719

Qy       6526 AAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACT 6585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACT 779

Qy       6586 GGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAAT 6645
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        780 GGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACANACTATCTCAAGAACTAAAT 839

Qy       6646 GGGGACAT-GGAACGGGTCGCCGACTCCCTGG-TCACCTTGCAAGATCAACTTAACTCCC 6703
             |||||||| ||||||| ||||||||| ||||| ||||||||||||||||||||||  ||
Db        840 GGGGACATGGGAACGGTTCGCCGACTTCCTGGTTCACCTTGCAAGATCAACTTAACCTCC 899

Qy       6704 TAGCAGCAGTAG-TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAG 6758
             ||||||||| || ||||||:|||||||||||  |||||||:|||| |||||| : |
Db        900 TAGCAGCAGNAGTTCCTTCAAAATCGAAGAGGCTTAGACTTGCTACCCGCTGTAGG 955
```

RESULT 11
BX439636/c
LOCUS BX439636 1019 bp mRNA linear EST 04-MAY-2004
DEFINITION BX439636 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24 3-PRIME, mRNA sequence.
ACCESSION BX439636
VERSION BX439636.2 GI:47000005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1019)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30771765.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE012DE12NP1&c=4215.r.
FEATURES Location/Qualifiers
source 1. .1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Query Match 11.8%; Score 886; DB 1; Length 1019;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 897; Conservative 45; Mismatches 32; Indels 2; Gaps 2;

```
Qy       6581 GTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAC 6640
             | : ::|||||||| : | :||||||||||||||||||||||||||||||||||||||
Db        975 GGWAYKGCATTGGCGTWTTAMCAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAC 916

Qy       6641 TAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTT-GCAAGATCAACTTAAC 6699
             |||||||||||||||||||||||||||||||||||||:|||||| ||||||||||||||||
Db        915 TAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGWCACCTTGGCAAGATCAACTTAAC 856

Qy       6700 TCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGA 6759
             |||||||||||||||||||||:|||||||||||||||||||:||||||||||||:|||
Db        855 TCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGA 796

Qy       6760 GGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATC 6819
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        795 GGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTACTATGTTAATCAATCCGGAATC 736

Qy       6820 GTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGA 6879
             |||||||||||||||:|||||||||:||||||||||||||||:|||||:|||||||||
Db        735 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGA 676

Qy       6880 AACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGA 6939
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        675 AACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGA 616
```

```
Qy       6940 CCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTT 6999
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        615 CCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTT 556

Qy       7000 AACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCC 7059
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        555 AACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCC 496

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        495 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAAT 436

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             ||||||||||||||||||||||||||||||||||||||:||||||| ||||||||||||||
Db        435 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGMACAACCTATACTACGCCCCAAT 376

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCCAACAGCACTTAGGTTTTCC 7239
             ||||||||||||||||||||||||:|||||||||||||||| ||||||||||||||||||
Db        375 TCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCCACAGCACTTAGGTTTTCC 316

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||: ||:|||||||:|||||||||
Db        315 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATYCCCYAGGCTGACTAAGAATCC 256

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||:|
Db        255 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCT 196

Qy       7360 CACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAG 7419
             |||||||||||||||||||||| |||||||||||||| ||||||||:||||||||::::
Db        195 CACACCTGACCAATCAGAGAGCTKACTAAAATGCTAACTAGGCAAAKACAGGAGGTWWWK 136

Qy       7420 AAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAA 7479
             ||||: ::|| ||:||||||:|||||||||||||||||||:||:|||||||||||||||
Db        135 TAATAKNVRATTATTTATTGCCTGAGAGCACAGCAGGAGGGACWATKATCGGGATATAAA 76

Qy       7480 CCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGC 7539
             :||||||:|||||||||||||| | ||||:::|||   ||:|: |  ||||| |||| :
Db         75 YCCAAGTCTTCGAGCCGGCAACAGAAACCYYYTTT-TTTCYCYCCTTTTTGTNTGGGNDT 17

Qy       7540 TCTGTTTTCATGCTAT 7555
             | | |||   | | |
Db         16 TTTTTTTNNTTTATTT 1
```

```
RESULT 12
BX378304
LOCUS       BX378304               1014 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  BX378304 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI018YH07 5-PRIME, mRNA sequence.
ACCESSION   BX378304
VERSION     BX378304.2  GI:46569716
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1014)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30443114.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI018CD04QP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .1014
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI018YH07"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
  Query Match            11.6%;  Score 867.4;  DB 1;  Length 1014;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  886;  Conservative   25;  Mismatches   24;  Indels    2;  Gaps    2;

Qy         80 TGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTC 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTC 60

Qy        140 TTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANA 199
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 TTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGA 120

Qy        200 CCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAG 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAG 180

Qy        260 CGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTA 319
              |||:|||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        181 CGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTA 240

Qy        320 AGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTC 379
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 AGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTC 300

Qy        380 TGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCC 439
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCC 360

Qy        440 TTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTG 499
              ||||||||||||||:||| ||||||||||||||||||||:|||||||||||||||||||
Db        361 TTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTG 419

Qy        500 GAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAA 559
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        420 GAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAA 479

Qy        560 GGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATGGGAAACGTTC 619
              ||||||||||||:||||||||||:||||:::|||||||||||:||||||||||||||||
Db        480 GGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATGGGAAACGTTC 539

Qy        620 CCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTC 679
              |||||||||||||||||||||||||||||||||||||:||||||||:||||||||||||
Db        540 CCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTC 599

Qy        680 AGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGA 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 AGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGA 659

Qy        740 AGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGT 799
              ||:||||||||||:|::|    | |   | |        ||||||||||||||||||||||
Db        660 AGWATAAATTATAAMAMMAAAAAAAAAAAAAAAAAAAAAAAATGTAGAAAAGGCAAATGGAGT 719

Qy        800 GAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAG 859
              |||||||||||||||||| :||:||||||||||||||||||||||||||||||||||||
Db        720 GAAGTGCCATAAGTACAAAAWTTMTTTTCATTAAGAGACAACTCACAATTATGTAAAAAG 779

Qy        860 TGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGAC 919
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         780 TGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGAC 839

Qy         920 TCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAA 979
               ||||||||||::|||||||||||||||||||||:||||||||||||||||||||:||
Db         840 TCCTTCCCCAACTAATAAGGACCCCCCTTCAACCMAAATGGTCCAAAAGGAGATAGAMAA 899

Qy         980 AAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCC 1016
               |||||||||:|||||| ||||||||||:|    | |||
Db         900 AAGGGTAAAMAGTGAA-CAAAGAGTGCMATATTCCCC 935


RESULT 13
BX347314
LOCUS       BX347314                924 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX347314 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI044YK06 5-PRIME, mRNA sequence.
ACCESSION   BX347314
VERSION     BX347314.1  GI:30365192
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 924)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAA016ZH06_CS01489_1&c=4215.r
            .

FEATURES             Location/Qualifiers
     source          1. .924
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI044YK06"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             11.2%;  Score 837.3;  DB 1;  Length 924;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  891;  Conservative   11;  Mismatches   16;  Indels    5;  Gaps    5;

Qy       5467 AGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGC 5526
              ||| ||||| |||||||| | || ||||||||||||||||||||||||||||| |||||||
Db          1 AGGCACCTGGAAATCCAGNAAAC-ACGCTAGCTATTCCTGTGAACCTCTAGA-GATTTGC 58

Qy       5527 GCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSC 5586
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||:|
Db         59 GCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCC 117

Qy       5587 CTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCC 5646
              |||||||||||||||||||:||:||||:|||:||||||:|||||||||||||||||||
Db        118 CTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCC 177

Qy       5647 CCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAG 5706
              |||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||

Db        178 CCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAG 237

Qy       5707 CGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTC 5766
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        238 CGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTC 297

Qy       5767 ACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCA 5826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCA 357

Qy       5827 AATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTC 5886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTC 417

Qy       5887 ACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGAT 5946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 ACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGAT 477

Qy       5947 CAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACC 6006
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        478 CAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACC 537

Qy       6007 TCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACT 6066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 TCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACT 597

Qy       6067 CGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 6126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 657

Qy       6127 CCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTA 6186
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        658 CCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA-GCCATATGTTTCAATCCCTGTA 716

Qy       6187 CCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT 6246
              |||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||
Db        717 CCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTNCGTTNTAGTAGGACCT 776

Qy       6247 CTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAAT 6306
              ||||||| | ||:||| ||||||||||||||||||||||||||||||||||||||||| ||
Db        777 CTTGTTTTCCATCTGGGAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCCAT 836

Qy       6307 ACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTC 6366
              |||||||||||||||| | || |||||||| |  ||||||||||||||| |||||||
Db        837 ACTACATACACAACCAAATTCCCATGCATCAAGGTGGTAACTCCTCCCACAC-AATAGTC 895

Qy       6367 TGCCTACCCTCAGGAATATTTTT 6389
              |||||||||||||||||||||||
Db        896 TGCCTACCCTCAGGAATATTTTT 918


RESULT 14
BX389656
LOCUS       BX389656                931 bp    mRNA    linear   EST 29-APR-2004
DEFINITION  BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION   BX389656
VERSION     BX389656.1  GI:30462930
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
```

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAF027ZB05_AF02531_1&c=4215.r

```
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             11.0%;  Score 824.3;  DB 1;  Length 931;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  899;  Conservative    8;  Mismatches   12;  Indels    7;  Gaps    7;

Qy       5848 AAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTC 5907
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AAAAAGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTC 61

Qy       5908 ACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTA 5967
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         62 ACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATGTA 121

Qy       5968 AAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTA 6027
             ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        122 AAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTA 181

Qy       6028 GATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAAT 6087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAAT 241

Qy       6088 ACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGC 6147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGC 301

Qy       6148 CTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTC 6207
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        302 CTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTC 361

Qy       6208 AGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATA 6267
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        362 AGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA 421

Qy       6268 ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCC 6327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCC 481

Qy       6328 CAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTT 6387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTT 541

Qy       6388 TTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTC 6447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTC 601

Qy       6448 CTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTC 6507
             ||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        602 CTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTC 661

Qy       6508 ATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTG 6567
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        662 ATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGNTATAGGAGCAGGAGTG 721

Qy       6568 CTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAA 6627
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |
Db        722 CTAGGTGCACTAGGTACTGGCATTGGCGG-ATCACAACCTCTACTCAGTTCTACTACANA 780

Qy       6628 CTATCTCAAGAACTAAATGGGGAC-ATGGAACGGGTCGCCGACT-CCCTGGTCACCTTGC 6685
             ||||||||||||||||||||||||| ||| |||||||||||||| ||||||||||||
Db        781 CTATCTCAAGAACTAAATGGGGACAATGTAACGGGTCGCCGACTCCCCTGGTCACCTTCG 840

Qy       6686 AAGATCAACTTAAC-TCCCTAGCAGCAGTA-GTCCTTCRAAATCGAAG-AGCTTTAGACT 6742
             |||||||||||| | |||||||||||||| | |||||||:||||||||| |||||||| ||
Db        841 CAGATCAACTTACCTTCCCTAGCAGCAGGAGGTCCTTCAAAATCGAAGAAGCTTTAGCCT 900

Qy       6743 YGCTAA-CCGCTGARAGAGGGGGAAC 6767
              :||||| |||||| :||||||||||
Db        901 TGCTAACCCGCTGCAAGAGGGGGAAC 926


RESULT 15
BX408752
LOCUS       BX408752                903 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX408752 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YI20
            5-PRIME, mRNA sequence.
ACCESSION   BX408752
VERSION     BX408752.2  GI:46931100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 903)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30648825.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZH03_CS00697_2&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .903
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE013YI20"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             11.0%;  Score 823;  DB 1;  Length 903;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches  869;  Conservative   12;  Mismatches   22;  Indels    6;  Gaps    4;

Qy       5531 GCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCC 5590
             ||||||  |||||||||||| || ||   |     ||     ||  ||||||||:|||||
Db          1 GCTCTTTCAACAACAACCAGNAGNAAGTAA---CTAAATNATAAATCCCCATGGCCCTCC 57

Qy       5591 CTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTC 5650
             ||||  ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db         58 CTTA-NATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTC 116
```

```
Qy       5651 CATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTC 5710
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        117 CATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTC 176

Qy       5711 CCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTG 5770
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        177 CCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTG 236

Qy       5771 CCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATA 5830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATA 296

Qy       5831 CTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTG 5890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 CTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTG 356

Qy       5891 TCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGG 5950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGG 416

Qy       5951 CAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTA 6010
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        417 CAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTA 476

Qy       6011 GCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCC 6070
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 GCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCC 536

Qy       6071 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTA 6130
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTA 596

Qy       6131 CTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTG 6190
             ||||||||||||||||||| |||||||||||:||||||||||||||||||||||||||||||
Db        597 CTAACTGTTGGATATGCCTNCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTG 656

Qy       6191 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 6250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 716

Qy       6251 TTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 6310
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 776

Qy       6311 CATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 6370
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 CATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 836

Qy       6371 TACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTT 6430
             |||||||||| |||||||||||||| |||||||||||||||   |||||| |||||||||
Db        837 TACCCTCAGG-ATATTTTTTGTCTG-GGTACCTCAGCCTATTCGTGTTTGGATGGCTCTT 894

Qy       6431 CAGAATCTA 6439
             |||||||||
Db        895 CAGAATCTA 903
```

```
RESULT 16
BX326647/c
LOCUS       BX326647                872 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX326647 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION   BX326647
VERSION     BX326647.1  GI:30334578
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 872)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAD010ZC11_AD00948_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
  Query Match             10.9%;  Score 813.5;  DB 1;  Length 872;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  844;  Conservative    6;  Mismatches   22;  Indels    3;  Gaps    2;

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             |||||||||||||| |||||||| |||||||||||||||||||||||  || ||||
Db        872 ATCCTAGTTGTCCTGAAGGACTTGAAGTCACTGTCTGTTGGACTTACTTCCACCCAACTT 813

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             |       |||||||||||||||||||||||||||||| |||||||||| ||||||||||
Db        812 GGTATGGCTATGGGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATGT-AAAGAAGTAA 754

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAN 694

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 634

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 574

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 514

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        513 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 454

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 394

Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 334

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 274

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
```

```
Db        273 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 214

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
Db        213 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 154

Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
Db        153 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 94

Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
Db         93 AACTAAATGGGGACATGGAACGGGTCGCCGACTCC--TGGTCACCTGCAAGATCAACTTA 36

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGA 6732
Db         35 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGA 1


RESULT 17
BX408733
LOCUS       BX408733                883 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  BX408733 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
            5-PRIME, mRNA sequence.
ACCESSION   BX408733
VERSION     BX408733.1  GI:30635957
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZE02_CS00686_1&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE012YJ24"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             10.8%;  Score 807;  DB 1;  Length 883;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  857;  Conservative   13;  Mismatches    9;  Indels    6;  Gaps    5;

Qy       5609 KTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 5668
Db          1 TTACTGTTCTTTTACCCTCTTT--ACTCTACTGCACCCCCTCCATGCCGCTGTATGACCA 58

Qy       5669 GTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 5728
Db         59 GTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 118

Qy       5729 CATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 5788
Db        119 CATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 178

Qy       5789 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 5848
Db        179 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 238

Qy       5849 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 5908
Db        239 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 298

Qy       5909 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 5968
Db        299 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 358

Qy       5969 AAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 6028
Db        359 AAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 418

Qy       6029 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 6088
Db        419 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 478

Qy       6089 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 6148
Db        479 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 538

Qy       6149 TCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 6208
Db        539 TCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 598

Qy       6209 GCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAA 6268
Db        599 GCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAA 658

Qy       6269 CCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCC 6328
Db        659 CCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTNCC 718

Qy       6329 AATGCATCA-GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTT 6387
Db        719 AATGCATCAGGGTGGGTAACTCCTTCCACACAAATAGTCTGCCTACCCTCA-GAATATTT 777

Qy       6388 TTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTC-TTCAGAATCTATGTGCTT 6446
Db        778 TTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTTCAGAATCTATGTGCTT 837

Qy       6447 CCTCTCATTCTTAGTGCCCCC-YATGRCCATCTACACTGAACAAG 6490
Db        838 CCTCTCATTCTTAGTGCCCCCTTATGACATTCTACACTGAACAAG 882


RESULT 18
BX380176/c
LOCUS       BX380176                966 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX380176 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI044YK06 3-PRIME, mRNA sequence.
ACCESSION   BX380176
VERSION     BX380176.2  GI:46834414
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 966)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
```

```
COMMENT     On May 8, 2003 this sequence version replaced gi:30460243.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI044BF03NP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .966
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI044YK06"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.8%;  Score 806.6;  DB 1;  Length 966;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  848;  Conservative   55;  Mismatches   54;  Indels    6;  Gaps    4;

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              | :|||||||| |||||  |  :||||||::|    |||:|:|||||||||||||||
Db        960 GGMTTGGCGGTTTCACA--CCTYACTCAGTTYWATACAAAAYTWTCTCAAGAACTAAAT- 904

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              ||||||:|||||||||||||: | ||||||||||||||||||||||||||||||||||||
Db        903 GGGACAKGGAACGGGTCGCSAATTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 844

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              | ||:||||||||||:|||||||||||||||:|||:||||:|||||||:||||||||||
Db        843 CGGCRGTAGTCCTTCAAAATCGAAGAGCTTTARACTTGCTAMCCGCTGAAAGAGGGGGAA 784

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 724

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:||||||||||||||||||:||||||:||||||||||||||||
Db        723 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTG 664

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        663 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 604

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        603 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 544

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
              |||||||||||||||||||:|||||||||||||||||||:||||||||||:|||||
Db        543 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCARGATGCAGTCCARGACTA 484

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
               |:|||||||| | |||||||||||||||:| |||| ||:|||||||||||||||||||
Db        483 GGDTCTACCGCGGCCCCCTGGACCGGCCTGCTSGCCCCCGMTCTGATGTTAATGACATCA 424

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
              ||||||||||||||| || |||||||||||||: |||||||||||||||||||| ||:|
Db        423 AAGGCACCCCTCCTGGGGCAATCTCAGCTGCACMCCCTCTACTACGCCCCAATTCGGCRG 364

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
              |||||||||:| |||||:|||||||||||||||| ||:||||||:|||||||||||||||
Db        363 GAAGCAGTTRGGGCGGTCGTCGGCCAACCTCCCCGACRGCACTTRGGTTTTCCTGTTGAG 304

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
              |||||||||||||||||:||||| ||||||||||||||||||:|||||||||:||:|||
Db        303 ATGGGGGACTGAGAGACMGGACTGGCTGGATTTCCTAGGCTGACTAAGAATCCCTARGCC 244

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
              |:|:|||||||||||||||||||||||||||||:|||||||||| |||| |:|||| |||
Db        243 TRGCTGGGAAGGTGACCACATCCACCTTTAAACMCGGGGCTTGCCACTTSGCTCACCCCT 184

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
              | ||||||:| | |||||||||||||||||:|: |||   | | :: ::    |   | ||
Db        183 GCCCAATCRGGGGGCTCACTAAAATGCTAMTYCGGCCCCGSCSSSSSVCSSSGCCCTCGC 124

Qy       7427 -CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAG 7485
               |: :|:|:|||||:|||||||||||:|||||||||||||||||||||||||||||:|
Db        123 GCMSYCMTCTATTGCCTGAGAGCACAGCMGGAGGGACAATGATCGGGATATAAACCCAMG 64

Qy       7486 TYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTT--GTATGGGAGCTCTG 7543
              |:|||| |||||| |||| ||||||||||||||||||||||     ||||||||||
Db         63 TCTTCGCGCCGGCCCCGGCCCCCCCCCTTTGGGTCCCCTCCCTTTCGTNNTGGGAGCTCTG 4

Qy       7544 TTT 7546
              | |
Db           3 TNT 1


RESULT 19
BX347619/c
LOCUS       BX347619                 877 bp    mRNA    linear   EST 23-APR-2004
DEFINITION  BX347619 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 3-PRIME, mRNA sequence.
ACCESSION   BX347619
VERSION     BX347619.1  GI:30379218
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAD009ZB06_AD00811_1&c=4215.r
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YJ18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.6%;  Score 791.2;  DB 1;  Length 877;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  832;  Conservative   12;  Mismatches   14;  Indels    4;  Gaps    4;
```

```
Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAA-CCCTACTAACTGTTGG 6141
Db        877 TAAAACCAACCCTCACTGGGCTCCATGAGTNCTCGGGCCAAAAACCCTACTAACTGTGGG 818

Qy       6142 ATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAAC 6201
Db        817 ATATGCCTCCNCCTGAACTTCAGGGCATATGTTTCAATCCCTGTACCTGAACAATGGAAC 758

Qy       6202 AACTTCAGCACAGA-AATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATST 6260
Db        757 ACTTTCAGCACAGANAATAAACACCACTNCCGTTTTAGTAGGACCTCTTGTTTCCAATCT 698

Qy       6261 GGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAAC 6320
Db        697 GGAAATAACCCATACCTCAAACCTCACCTGTGT-AAATTTAGCAATACTACATACACAAC 639

Qy       6321 CAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGG 6380
Db        638 CAACTCNCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGG 579

Qy       6381 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTAT 6440
Db        578 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTAT 519

Qy       6441 GTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAG 6500
Db        518 GTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAG 459

Qy       6501 TTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGC 6560
Db        458 TTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGC 399

Qy       6561 AGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTA 6620
Db        398 AGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTA 339

Qy       6621 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCAC 6680
Db        338 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCAC 279

Qy       6681 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGA 6740
Db        278 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGA 219

Qy       6741 CTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 6800
Db        218 CTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 159

Qy       6801 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACG 6860
Db        158 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACG 99

Qy       6861 TAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTG 6920
Db          98 TAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCC-G 40

Qy       6921 GATTCTCCCCTTCTTAGGACCT 6942
Db          39 GATTCTCCCCTTCTTAGGACCT 18
```

RESULT 20
BX408734
LOCUS BX408734 921 bp mRNA linear EST 30-APR-2004
DEFINITION BX408734 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24 5-PRIME, mRNA sequence.
ACCESSION BX408734
VERSION BX408734.1 GI:30635959
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZE02_CS00686_2&c=4215.r

FEATURES Location/Qualifiers
source 1. .921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Query Match 10.4%; Score 779.701; DB 1; Length 921;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 853; Conservative 13; Mismatches 47; Indels 7; Gaps 5;

```
Qy       5609 KTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 5668
Db          1 TTTCTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 60

Qy       5669 GTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 5728
Db         61 GTAGCTCCCCTTACCAANAGTTTCTATGGAGAATGCAGCGTCCCGG-AATATTGATGCCC 119

Qy       5729 CATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 5788
Db        120 CATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 179

Qy       5789 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 5848
Db        180 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 239

Qy       5849 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 5908
Db        240 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 299

Qy       5909 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 5968
Db        300 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 359

Qy       5969 AAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 6028
Db        360 AAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 419

Qy       6029 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 6088
Db        420 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 479

Qy       6089 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 6148
Db        480 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 539

Qy       6149 TCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 6208
```

```
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 TCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 599

Qy       6209 GCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAA 6268
              |||||||||| ||||||||||| ||||||||||||||||||||||||:|||||||||
Db        600 GCACAGAAATTAACACCACTTCCCGTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAA 659

Qy       6269 CCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCC 6328
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        660 CCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACC-ACTTCC 718

Qy       6329 AATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTT 6388
              |||| || |||||||||||||||| || |||||||||||||||||||||| | |||||||
Db        719 AATGGATTAGGTGGGTAACTCCTTCCCCACAAATAGTCTGCCTACCCTCAAGGATATTTT 778

Qy       6389 TTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCC 6448
              ||| |||||| |||||||||||||| | ||| |||| ||||||||| || | ||||||
Db        779 TTGGCTGTGGGACCTCAGCCTATCGGCGGTTGGATGGGTCTTCAGAA-ATAAGGGCTTCC 837

Qy       6449 TCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCA 6508
              |||||  ||  |  |||| :  : |  || ||||| |  ||| ||||||| ||   ||
Db        838 CCTCATCCTAGGGCCCCCAT---GACCCCTTCACTGGAACAGA-TTATACAATTTGTTCT 893

Qy       6509 TATCTAAGCCCCGCAACAAA 6528
                ||||  |||| | |||||
Db        894 ATACTAACCCCCACCACAAA 913


RESULT 21
BX389657
LOCUS       BX389657                995 bp    mRNA    linear   EST 29-APR-2004
DEFINITION  BX389657 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION   BX389657
VERSION     BX389657.2  GI:46875146
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 995)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30462931.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAF027ZB05_AF02531_2&c=4215.r
            .

FEATURES             Location/Qualifiers
     source          1. .995
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.3%;  Score 774.602;  DB 1;  Length 995;
    Best Local Similarity   91.0%;  Pred. No. 0;
    Matches  866;  Conservative    7;  Mismatches   67;  Indels   12;  Gaps    5;

Qy       5846 GAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACT 5905
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 GAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGT-ACTGTCTGTTGGACTTACT 59

Qy       5906 TCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATG 5965
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         60 TCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATG 119

Qy       5966 TAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGAC 6025
              ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        120 TAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGAC 179

Qy       6026 TAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTA 6085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTA 239

Qy       6086 ATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATAT 6145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 ATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATAT 299

Qy       6146 GCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACT 6205
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACT 359

Qy       6206 TCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAA 6265
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        360 TCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAA 419

Qy       6266 TAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACT 6325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACT 479

Qy       6326 CCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATAT 6385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATAT 539

Qy       6386 TTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCT 6445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 TTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCT 599

Qy       6446 TCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATG 6505
              | |    |||||||||||||||:|||:|||||||||||||||||||||||||||||||||||||
Db        600 TGCCTCTATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATG 659

Qy       6506 TCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAG 6565
              |||||||||  |||||||||||||||||||||||||||||||  |||||||     |||
Db        660 TCATATCTAAAGCCCGCAACAAAAGAGTACCCATTCTTCCTTTTTGTATAGGAAGCAGAG 719

Qy       6566 TGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACA 6625
              |||||||||||||||| | ||| |||||| || || ||| ||| || ||| ||||||||||
Db        720 TGCTAGGTGCACTAGATCCTGCCATTGGTGGAATAACATCCTTTATTCATTTCTACTACA 779

Qy       6626 AACTATCTCAA-GAACTAAATGGGGACATGGAACGGGTCGCC---GACTCCCTGGTCACC 6681
              |||||| | || || ||||||||| |||||||||||||| | |      | ||  |    |
Db        780 AACTATTTNAAGGACCTAAATGGGAACATGGAACGGGTTGGCCGACTCCCCTGGTAACCT 839

Qy       6682 TTGCAAGATCAACTTAACTCCC----TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              | | ||||||||||||||| |     ||| | ||       |  :||    |||  ||||
Db        840 TGGAAAGATCAACTTAACTTGCCTAACAGCCGGAGGTTCTTTCAAAATCGCAAGGCCTTT 899

Qy       6738 AGACTYGCTAACCGCTG---ARAGAGGGGGAACCTGTTTATTTTTAGGGGAA 6786
              |||||  ||||||||||    |:|| |||| | | ||||||||||||||||
Db        900 AGACTGTCTAACCGCTGTAAAGAGGGGGGAACCTTGTTTATTTTTAGGGGAA 951


RESULT 22
BX388766/c
LOCUS       BX388766                836 bp    mRNA    linear   EST 29-APR-2004
DEFINITION  BX388766 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI044YK06 3-PRIME, mRNA sequence.
```

```
ACCESSION   BX388766
VERSION     BX388766.2  GI:46877917
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30462383.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AU012ZB04_U01082_1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .836
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI044YK06"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.3%;  Score 772;  DB 1;  Length 836;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  799;  Conservative   12;  Mismatches   22;  Indels    2;  Gaps    2;

Qy       6144 ATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAA 6203
             |||||||||||||||||  |:|||| ||||| |||||||||   ||| | ||||||  ||
Db        836 ATGCCTCCCCCTGAACTCAAGGCCAAATGTTCCAATCCCTGGTCCTCACCAATGGCCCAC 777

Qy       6204 CTTCAGCACAGAAA-TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
             ||||||| |||||| || |||||||| |||||| |||||||||||||||||||||:|||
Db        776 CTTCAGCCCAGAAATTACACACCACTCCCGTTTAAGTAGGACCTCTTGTTTCCAATCTGG 717

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
             |||||||||| |||||||||||||||||||||||| |||||| |||||||||| |||
Db        716 AAATAACCCATTCCTCAAACCTCACCTGTGTAAAATTAAGCAATCCTACATACACACCCA 657

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        656 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 597

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 537

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
             |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        536 GCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTT 477

Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 417

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 357

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 297

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        296 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACT 237

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATG 6802
             :|||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 TGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATG 177

Qy       6803 TTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTA 6862
             ||||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||
Db        176 TTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTA 117

Qy       6863 KAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGA 6922
             :||||||:||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        116 GAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGA 57

Qy       6923 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACC 6977
             ||||||||||||||||||||||||||||||||||||||||| ||||||||| ||| || ||
Db          56 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA-TGCTACTCCACTTGGGCCC 3
```

```
RESULT 23
DA834147
LOCUS       DA834147                808 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834147 PLACE3 Homo sapiens cDNA clone PLACE3000261 5', mRNA
            sequence.
ACCESSION   DA834147
VERSION     DA834147.1  GI:83037485
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE3000261"
                     /tissue_type="placenta"
                     /clone_lib="PLACE3"
                     /note="Vector: pME18SFL3"

  Query Match             10.3%;  Score 768.6;  DB 1;  Length 808;
```

```
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  783;  Conservative   14;  Mismatches   10;  Indels    2;  Gaps    2;

Qy         81 GTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCT 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCT 60

Qy        141 TACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANAC 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db         61 TACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGAC 120

Qy        201 CTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGC 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGC 180

Qy        261 GARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAA 320
              ||:|||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 GAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAA 240

Qy        321 GTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCT 380
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        241 GTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCT 300

Qy        381 GTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCT 440
              ||||||||||||||||||:|||| ||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCCACGGCTTCTAATAGAACTTTAACACTTACCACATGGCCCAAGATTCCATTCCT 360

Qy        441 TGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGG 500
              |||||||||||||||:||| ||||||||||||||||||||:|||||||||||||||||||||
Db        361 TGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGG 419

Qy        501 AAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAG 560
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAG 479

Qy        561 GACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATGGGAAACGTTCC 620
              ||||||||||:||||||||||:||||:::||||||||||:|||||||||||||||||||||
Db        480 GACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATGGGAAACGTTCC 539

Qy        621 CCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCA 680
              ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||||
Db        540 CCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCA 599

Qy        681 GACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAA 740
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        600 GACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGNCTGGCACTCCTGAGGGAA 659

Qy        741 GTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTG 800
              |||||||||||||||||||||||||||||||:|||||||||||||||||| ||||| ||
Db        660 GTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCANATGGANTG 719

Qy        801 AAGTGCCATAAGTACAAACTTTCTTTTCATT-AAGAGACAACTCACAATTATGTAAAAAG 859
              || ||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        720 AAATGCCATAAGTACAAACTTTCTTTTCATTAAAGAGACAACTNACAATTATGTAAAAAG 779

Qy        860 TGTGATTTATGCCCTACAGGAAGCCTTCA 888
              ||||| |||||||||||||||||||||||
Db        780 TGTGAATTATGCCCTACAGGAAGCCTTCA 808


RESULT 24
BX347111/c
LOCUS       BX347111                844 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX347111 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YE21 3-PRIME, mRNA sequence.
ACCESSION   BX347111
VERSION     BX347111.1  GI:30375145
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 844)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            7108.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AU010ZB08_U0922_1&c=7108.r.
FEATURES             Location/Qualifiers
     source          1. .844
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YE21"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.2%;  Score 766.4;  DB 1;  Length 844;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  791;  Conservative   11;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       6082 TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAA-CCCTACTAACTGTTG 6140
              || ||| |||||||| ||||| ||||| |||||||||||||||| |||||||||||||| |
Db        837 TTAAATCCCACCCTCCCTGGGTTCCATTAGGTCTCGGCCCAAAACCCCTACTAACTGTGG 778

Qy       6141 GATATGCCT-CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGA 6199
              ||||||||| ||||||||| ||||:||||||||||||||||||||||||||||||||||||
Db        777 GATATGCCTCCCCCCTGAATTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGA 718

Qy       6200 ACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATS 6259
                |  |||||||||||||||||| ||||||||||||||||||||||||||||||||||:
Db        717 CAACTTTCAGCACAGAAATAAACCCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATC 658

Qy       6260 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 598

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 CCAACTCNCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 538

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 478

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        477 TGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACA 418

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 358

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 298

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 238
```

```
Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAG 6739
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        237 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAG 178

Qy       6740 ACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATT 6799
             |||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        177 ACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATT 118

Qy       6800 ATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAAC 6859
             ||||||||||||||||||||||||||||||||||:|||||||||:|||||||||||||||
Db        117 ATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAC 58

Qy       6860 GTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCC 6901
             |||:|||||||:|||||||||||||||||||  ||| |  ||
Db         57 GTAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGGCCGCCC 16


RESULT 25
BX367907
LOCUS       BX367907                828 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX367907 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION   BX367907
VERSION     BX367907.1  GI:30459586
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AT022ZG01_T02023_1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .828
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.2%;  Score 765.8;  DB 1;  Length 828;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches  788;  Conservative    6;  Mismatches    8;  Indels    2;  Gaps    2;

Qy       5690 TTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTA 5749
             ||   |||||||||||||||||| |||||||||||||||||||||||||||||||:||
Db          9 TTTCTTGGAGAATGCAGCGTCCCGG-AATATTGATGCCCCATCGTATAGGAGTCTTTCTA 67

Qy       5750 AGGGAACCCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCA 5809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 AGGGAACCCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCA 127

Qy       5810 CTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTC 5869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        128 CTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTC 187

Qy       5870 CTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATG 5929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATG 247

Qy       5930 GGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCA 5989
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        248 GGGGTGGAGTTCAAGATCAGGCAAGAAAAAAACATGTAAAAGAAGTAATCTCCCAACTCA 307

Qy       5990 CCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAA 6049
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAA 367

Qy       6050 CCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATG 6109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATG 427

Qy       6110 AGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCAT 6169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        428 AGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCAT 487

Qy       6170 ATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTT 6229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 ATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTT 547

Qy       6230 CCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCT 6289
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        548 CCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCT 607

Qy       6290 GTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTC 6349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTC 667

Qy       6350 CTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCT 6409
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||
Db        668 CTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGTTACCTCAACCT 727

Qy       6410 ATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYA 6469
             |||||||||||||||||| |||||||||||| |||||||||||||||||| ||||||:|
Db        728 ATCGTTGTTTGAATGGCTC-TCAGAATCTATGGGCTTCCTCTCATTCTTAGGGCCCCCTA 786

Qy       6470 TGRCCATCTACACTGAACAAGATT 6493
             ||:|||||||||||||||||||||
Db        787 TGACCATCTACACTGAACAAGATT 810


RESULT 26
BX368078/c
LOCUS       BX368078                846 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  BX368078 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION   BX368078
VERSION     BX368078.2  GI:46572908
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30447710.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
```

division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0AU012ZD05_U01092_1&c=4215.r.

```
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI051YM13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             10.1%;  Score 756.3;  DB 1;  Length 846;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  810;  Conservative   11;  Mismatches   15;  Indels    5;  Gaps    5;

Qy       6047 AAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCAC-CCTCACTGGGCTC 6105
              |||||  | |||||||||||||| || |||||||||||||||| ||  ||||||||||||
Db        846 AAACCTCCGGTACCCATACTCGGCTTGTAAGCCTATTTAATAACAAACCTCACTGGGCTC 787

Qy       6106 CATGAGGTCTCGGCCC-AAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA- 6163
                ||||||||||||| |||||||||||||||| | ||||||||||||||||||||||
Db        786 ATATAGGTCTCGGCCCAAAAACCCTACTAACTGTGGAATATGCCTCCCCCTGAACTTCAG 727

Qy       6164 RGCCATATGTTTCAAT-CCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAAC 6222
              :|||||||||||||| |||||||||| |||||||||||||||||||||||||||||||
Db        726 GGCCATATGTTTCAATNCCCTGTACCTGANCAATGGAACAACTTCAGCACAGAAATAAAC 667

Qy       6223 ACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAAC 6282
              |||||| ||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        666 ACCACTNCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAAC 607

Qy       6283 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 6342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 547

Qy       6343 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 6402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 487

Qy       6403 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 6462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 427

Qy       6463 CCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 6522
              |||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 367

Qy       6523 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 6582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 307

Qy       6583 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 6642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 247

Qy       6643 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 187

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
              ||||||||||||||||||:||||||||||||||||||:||||||||||:||||||
Db        186 CTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGG 127

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 6822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 67
```

```
Qy       6823 ACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAAC 6882
              |||||||||||:|||||||||:||||||||||||||||| :|||||||:||||||||||||
Db          66 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGT-GAGCAGAGGAGCTTCGAAAC 8

Qy       6883 A 6883
              |
Db           7 A 7
```

```
RESULT 27
DA833993
LOCUS       DA833993                820 bp    mRNA    linear   EST 15-NOV-2005
DEFINITION  DA833993 PLACE2 Homo sapiens cDNA clone PLACE2000344 5', mRNA
            sequence.
ACCESSION   DA833993
VERSION     DA833993.1  GI:82430788
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000344"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"

  Query Match             10.1%;  Score 753.3;  DB 1;  Length 820;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  771;  Conservative   31;  Mismatches   12;  Indels    3;  Gaps    3;

Qy       4132 AATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGC 4191
              ||||||:|:|||||||||||||||||||:|||:|||||||||||:||||||||||||
Db          1 AATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGC 60

Qy       4192 CATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYG 4251
              ||||||||:||||||:||||:||||||||||||:||||||||:||||||||||||:|
Db         61 CATTAGGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCG 120

Qy       4252 GGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGC 4311
              ||||||||| |||||||||:||||||||||:|||:|||:|||||||||:||||||||||
Db        121 GGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGAGAACTGCCAAGCAGATATTGAAGC 180

Qy       4312 MAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATA 4371
              :|||||||||||||||||||||||||||||||||||||||||||||| ||||  ||| |||||||
```

```
Db        181 CAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATA 240

Qy       4372 GGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAAC 4431

Db        241 GGGTAATCCCCTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAAC 300

Qy       4432 CTCACGAGG-CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTT 4490

Db        301 CTCACGAGGACAGTTTTCTCCCCTCGGGACGGCTAGCCACTGAAGAAGGGAAAATACTTT 360

Qy       4491 TGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCA 4550

Db        361 TGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCA 420

Qy       4551 TCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTA 4610

Db        421 TCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTA 480

Qy       4611 TCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAAATAATCCCCTGCCTYATCGCC 4670

Db        481 TCAAGCAGATAGTCAGGGCCTGTGTAGTGTGCCAGAGAAATAATCCCCTGCCTTATCGCC 540

Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTT 4730

Db        541 AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTT 600

Qy       4731 TACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCAC 4790

Db        601 TACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTTTCAC 660

Qy       4791 GGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACT 4850

Db        661 GGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGC-CT 719

Qy       4851 AGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCC 4910

Db        720 AGTTCATGAAATAATT-CCAGATTCGGACTTCCCCGANGCTTACAGAGTGACAATAGCCC 778

Qy       4911 TGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAG 4947

Db        779 TGCTTTTCAGGCCACAGTAACCCANGGAGTATNCCAG 815
```

RESULT 28
BX459153/c

```
LOCUS       BX459153                900 bp    mRNA    linear   EST 05-MAY-2004
DEFINITION  BX459153 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YI20
            3-PRIME, mRNA sequence.
ACCESSION   BX459153
VERSION     BX459153.2  GI:47051796
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 22, 2003 this sequence version replaced gi:31021096.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DE013BE10NP1&c=4215.r.
```

```
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE013YI20"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             10.0%;  Score 752.6;  DB 1;  Length 900;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  768;  Conservative   71;  Mismatches   59;  Indels    2;  Gaps    2;

Qy       6622 TACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACC 6681

Db        900 TACAAACTATCTCAAGAACTAAATGGGGACATGGAASGGGTCGCCGACTCCCTGGTCACC 841

Qy       6682 TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGAC 6741

Db        840 TTGCAAGATCAACTTAACGCCCTAGCAGCAGTAKTCCTTCAAAATCGAAGAGCTTTAGAC 781

Qy       6742 TYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTAT 6801

Db        780 TTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTACTAT 721

Qy       6802 GTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGT 6861

Db        720 GTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGT 661

Qy       6862 AKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGG 6921

Db        660 AGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGG 601

Qy       6922 ATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGT 6981

Db        600 ATTCTCCCCCTCCTASGACCTCTMGCAGCTATAATATTGCTACTCCTCTTTSGACCCTG- 542

Qy       6982 ATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATG 7041

Db        541 ATCTTTAACCTCCTTG-TAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACCACAAATG 483

Qy       7042 GAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGC 7101

Db        482 GAGCCCAAGATGCAGCCCAASACCAASACCTACCGCAGACCCCTGGACCSGCCCGCTMGC 423

Qy       7102 CCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAA 7161

Db        422 CCACSATCTGATGTCAATSACACCAAASGCACCCCCCCCMGASGAAATCCCAGCCGCACAA 363

Qy       7162 CCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCA 7221

Db        362 CCTCTACTACGCCCCAATTCAGCASGAAGCAGTTAGAGCGGTCSTCSCCCAACCCCCCCM 303

Qy       7222 ACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCC 7281

Db        302 ACAGCACTTASGTTTCCCTGTTSAGATGGGSCACTSAGASACAGSACTAGCTSSATCMCC 243

Qy       7282 TAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACAC 7341

Db        242 CMCCCCCACCMMCAACCCCCAACCCTMGCTGGGAAGGTGACCACACCCMCCCCCAMACAC 183

Qy       7342 GGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGG 7401

Db        182 GGSGCCCSCCACTTCGCTCACVCCTGSCCAATCAGAGAGCGSRSGGAAAGGCTRRCTAGG 123

Qy       7402 CAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGA 7461

Db        122 CAAAGACAGGAGGCARGGRRATMGCCMVCCATSCACTGCCTGAGAGCACRGCAGGAGGGA 63
```

```
Qy       7462 CAATGATCGGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCCTTTGGGTCCC 7521
              |:|||: ||||| :||||:||||||:: || |||||| ||||| |||||||:|||  |||
Db         62 CVATGMGCGGGAGRTAAASCCAAGTCBCCGCGCCGGCTACGGCKCCCCCCTKTGGTCCCC 3


RESULT 29
CZ457645
LOCUS       CZ457645                836 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF745b17TR Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_45b17, genomic survey sequence.
ACCESSION   CZ457645
VERSION     CZ457645.1  GI:77934335
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .836
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_45b17"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             9.8%;  Score 734.7;  DB 1;  Length 836;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  738;  Conservative   53;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:||||||||||: ||||| ||  |||||||||
Db          1 GAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACTGGTAGGGGCAGCAGTAGCA 60

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||| |||||| ||||||||||||||||||:|||||||||||||||||||||||||
Db         61 GTCTTAGTTTCTGAAACAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCT 120

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::|||||||||||||||||| ||||||||||||||||||||||: ||||
Db        121 CATGATGTGAACGGCATACTCACTGCTAAAGAAGACTTGTGGCTGTCAGACAACCATTTA 180

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||: |||| |||||||||||||: ||||||||||:||||  || ||||||||||
Db        181 CTTAAATAGCAGGTTCTATTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACT 240

Qy       3775 CTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAA 3834
              ||||||||||:| ||||||||||||||||||||||||||||:||:|||||||||||||||
Db        241 CTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAA 300

Qy       3835 RTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTATGAGRGTTCCYTTGACTGATCCY 3894
              :||||| |||||||||||||  ||||||||||||||:||||:|||||:||||||||||||:
Db        301 GTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCC 360

Qy       3895 GACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGG 3954
              |||| |||||||||||||||||||||||||||| | ||||||||||||| ||||||:|||
Db        361 GACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGG 419

Qy       3955 GTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAG 4014
              ||||||||| |||||||||||||||:|||||||||||| ||||||||||||||||||||
Db        420 GTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAG 479

Qy       4015 TGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAG 4074
              ||||:| ||:|||||||||||||||:|||:|:|||||||||||||||||||||:||||||
Db        480 TGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAG 539

Qy       4075 GGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAAT 4134
              ||:|||||||||| ||||||||:|:|||||||:|||||||| ||||||||||||::|||||
Db        540 GGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAAT 599

Qy       4135 ATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCAT 4194
              |||:|:||| |||||||||||||||:|||:|||||||||||: |||||||||||||||||
Db        600 ATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACCATCAGGAAGCCAT 659

Qy       4195 TAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGG 4254
              |||||:|||||||:||||:||||||||||||:||||||||:|||||||||||||: |||
Db        660 TAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGG 719

Qy       4255 TCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAA 4314
              |||||| |||| | ||:|||||||||||:|||: ||: ||||||| :||||||||||:||
Db        720 TCATCAGAAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAA 779

Qy       4315 AAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTAT 4370
              || ||| |||||||||||| ||||||||||||||||||||  ||   ||| ||||||
Db        780 AAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTAT 835


RESULT 30
DR004257
LOCUS       DR004257                770 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC114945 Human placenta, large insert, pCMV expression library Homo
            sapiens cDNA clone TC114945 5' similar to Homo sapiens endogenous
            retroviral family W, env(C7), member 1 (syncytin) (ERVWE1), mRNA
            sequence.
ACCESSION   DR004257
VERSION     DR004257.1  GI:66264130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
```

```
           6 Taft Ct. Suite 100
           Rockville, MD 20850
           Tel: (301) 340-3188
           http://www.origene.com
           Seq primer: pCMV6 5prime forward vector primer, OriGene
           Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC114945"
                     /tissue_type="Placenta"
                     /clone_lib="Human placenta, large insert, pCMV expression
                     library"
                     /note="Vector: pCMV6-XL4; Site_1: EcoR1; Site_2: Xho1/Sal1
                     compatible end ligatio; Oligo-dT primed reverse
                     transcription optimized for large and GC rich mRNA
                     transcripts, cDNA size selection, optimized ligation for
                     large inserts into mammalian expression vector, random
                     clones selected for end sequence verification of
                     full-length genes"

  Query Match             9.8%;  Score 733.6;  DB 1;  Length 770;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches  748;  Conservative   13;  Mismatches    3;  Indels    2;  Gaps    2;

Qy       5082 GTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATA 5141
             ||||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db          7 GTTGCCTATAGCCTTAAAAAAAATCTGCAACTTTCCCCAAAAAGCA-GACTTAGCCCATA 65

Qy       5142 CGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCC 5201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCC 125

Qy       5202 AACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAA 5261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 AACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAA 185

Qy       5262 GGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTC 5321
             ||||||||||||||||||||||||||||||||||||||::|||||||||||||||||||
Db        186 GGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTC 245

Qy       5322 AAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTC 5381
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTC 305

Qy       5382 ATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACT 5441
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACT 365

Qy       5442 TGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTAT 5501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 TGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTAT 425

Qy       5502 TCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAA 5561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 TCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAA 485

Qy       5562 CTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTY 5621
             ||||||||||||| |||||||:|||||||||||||||||:||:||||:|||:
Db        486 CTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTT 544

Qy       5622 ACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYA 5681
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        545 ACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTA 604

Qy       5682 CCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAG 5741
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAG 664

Qy       5742 TCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCA 5801
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        665 TCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCA 724

Qy       5802 CTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGA 5847
              |||||||||||||||||||||||  ||||||||||||||||||||
Db        725 CTCTGCCACTCTTTGCATGCATGNCAATACTCATTATTGGACAGGA 770


RESULT 31
DA833923
LOCUS       DA833923               769 bp    mRNA    linear   EST 15-NOV-2005
DEFINITION  DA833923 PLACE2 Homo sapiens cDNA clone PLACE2000161 5', mRNA
            sequence.
ACCESSION   DA833923
VERSION     DA833923.1  GI:82430486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000161"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"

  Query Match             9.8%;  Score 731.6;  DB 1;  Length 769;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  749;  Conservative   11;  Mismatches    8;  Indels    2;  Gaps    2;

Qy       4855 CATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCT 4914
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 CATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCTCTGCT 60

Qy       4915 TTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACAC 4974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACAC 120

Qy       4975 TGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGA 5034
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        121 TGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGA 180

Qy       5035 CATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCC 5094
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        181 CATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCC 240
```

```
Qy       5095 TTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTAT 5154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTAT 300

Qy       5155 GGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAG 5214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        301 GGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTTGTTGCAG 360

Qy       5215 ACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCT 5274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCT 420

Qy       5275 GAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCT 5334
             ||||||||||||||||||||||||||::||||||||||||||||||||||||||||:||
Db        421 GAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCT 480

Qy       5335 CTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCC 5394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        481 CTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCC 540

Qy       5395 CAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCT 5454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCT 600

Qy       5455 GGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTC 5514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTC 660

Qy       5515 TAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAA 5574
             |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        661 TAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGANGAAAGT-ACTAAAATCATAAA 719

Qy       5575 TCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACC 5624
             | |||||||:||| |||||||||| | ||||||:||:||  :|||:|||
Db        720 T-CCCCATGGGCCTTCCTTATCATAATNTTCTCTTTACTGGNCTTTTACC 768


RESULT 32
BX325902
LOCUS       BX325902                879 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX325902 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI018YH07 5-PRIME, mRNA sequence.
ACCESSION   BX325902
VERSION     BX325902.2  GI:46269730
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30334514.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AT019ZH04_T01760_1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI018YH07"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             9.7%;  Score 729.9;  DB 1;  Length 879;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  840;  Conservative   15;  Mismatches   23;  Indels   13;  Gaps   10;

Qy        695 AACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTG-AGGGAAGTATAAATTATAA 753
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          1 AACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGNAGGGAAGTATAAATTATAA 60

Qy        754 CACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGT 813
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db         61 CACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGT 120

Qy        814 ACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCC 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCC 180

Qy        874 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTA 933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::||
Db        181 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTA 240

Qy        934 ATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTG 300

Qy        994 AACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAA 1053
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        301 AACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAA 359

Qy       1054 TTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACA 1113
              |||||||||||||||||||||||:||||::|||||||||||||||||||||||||||||||
Db        360 TTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACA 419

Qy       1114 GACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGA 1173
              |||||||||||||||||||||||:|||||||||||||||:|||||||||||||||||||||
Db        420 GACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 479

Qy       1174 CAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCC 1233
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        480 CAATTCTTTGATCTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCC 538

Qy       1234 AAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAG 1293
              |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        539 AAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAG 598

Qy       1294 TCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCA 1353
              |||||||||||     |||||||||||| |||||||| ||||||||||||||||||||||||
Db        599 TCAGGTCAATG--ATAGGATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCA 655

Qy       1354 RGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGT 1413
              :|||||||||||||:|||||||||||    ||| | |     |||| ||||||||||
Db        656 GGCAGTTCCCAGTCTAGACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGT 711

Qy       1414 GCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCT 1473
              ||||||||||||||||||||||||||:|||    | | ||||||:|||| ||:|||
Db        712 GCTGCAGACATTTGCTAACTTGTGTGCTAGAAGGGACTAAGGAAACTAGGAAG-AAGTCT 770

Qy       1474 AYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTT 1533
              |:|||||||||||||||||||||||| |||||||| |||||||| ||||||||||||||
Db        771 ATGAATTACTCAATGATGTCCACCATTACACAGGGAAAGGGAAG-AAATCCTACTGCCTT 829

Qy       1534 TCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGA 1584
              ||| || ||||||| |||||||||||||||||||||| |||||||| |||
Db        830 TCT-GAAAGACTAAAGGAGGCATTGAGGAAGCGTGCCTTTCTGTCACTTGA 879

RESULT 33
```

```
BX357208/c
LOCUS       BX357208                758 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX357208 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 3-PRIME, mRNA sequence.
ACCESSION   BX357208
VERSION     BX357208.2  GI:46305595
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30376125.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI022DE09NP1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YJ18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             9.7%;  Score 724.6;  DB 1;  Length 758;
  Best Local Similarity   95.8%;  Pred. No. 0;
  Matches  712;  Conservative   26;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       6794 GTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAA 6853
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||
Db        758 GTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAA 699

Qy       6854 TACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGA 6913
             |||||||||:||||||:||||||||||||||||||||||||||||||||||:|||||
Db        698 TACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGA 639

Qy       6914 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 6973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 579

Qy       6974 GACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAAC 7033
             |||||||||||||:||||||||||||||||||||||||||||||||||||||:||||
Db        578 GACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAAC 519

Qy       7034 TACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGC 7093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 TACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGC 459

Qy       7094 CTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAG 7153
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAG 399

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        398 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAA 339

Qy       7214 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 7273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 279

Qy       7274 GGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCT 7333
             ||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||
Db        278 GGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCT 219

Qy       7334 TTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 7393
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        218 TTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 159

Qy       7394 TAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGC 7453
             ||||||||||||||||||||||||||||||:||||||||||:||||:|:::||||||||||
Db        158 TAATTAGGCAAAGACAGGAGGTAAAGAAMTAGCCAATCATCTATTSCCYSAGAGCACAGC 99

Qy       7454 AGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTT 7513
             ||||||||||||||||||||||||||||||:::::  || ||||||:||||:||||||:
Db         98 AGGAGGGACAATGATCGGGATATAAACCCMMSYCCCCGCGCCGGCCACGGCMACCCCCYC 39

Qy       7514 TGGGTCCCCTCCCTTTGTATGGG 7536
             |||||||||||||||||||||||
Db         38 TGGGTCCCCTCCCTTTGTATGGG 16


RESULT 34
BX409328
LOCUS       BX409328                924 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX409328 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YI20
            5-PRIME, mRNA sequence.
ACCESSION   BX409328
VERSION     BX409328.2  GI:46932867
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 924)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30652931.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAE012ZF02_AE01071_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .924
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE013YI20"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             9.7%;  Score 723.302;  DB 1;  Length 924;
  Best Local Similarity   89.8%;  Pred. No. 0;
```

```
  Matches  832;  Conservative   10;  Mismatches   75;  Indels    9;  Gaps    7;

Qy       5531 GCTCTTCAAACAACAACCAGGAGGAAA-GTAACTAAAATCATAAATCCCCCATGGSCCTC 5589
              ||||||  |||||||||||  |||||||  ||||||||||||||||||||||    ||:||||
Db             1 GCTCTTTAAACAACAACCCGGAGGAAAGGTAACTAAAATCATAAATCCCCATGGGCCCTC 60

Qy       5590 CCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCT 5649
              |||||  ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db            61 CCTTA-NATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCT 119

Qy       5650 CCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGT 5709
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db           120 CCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGT 179

Qy       5710 CCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACT 5769
              |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db           180 CCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACT 239

Qy       5770 GCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAAT 5829
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           240 GCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAAT 299

Qy       5830 ACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACT 5889
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           300 ACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACT 359

Qy       5890 GTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAG 5949
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           360 GTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAG 419

Qy       5950 GCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCT 6009
              |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db           420 GCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCT 479

Qy       6010 AGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGC 6069
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           480 AGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGC 539

Qy       6070 CTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCT 6129
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db           540 CTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCANAACCCT 599

Qy       6130 ACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCT 6189
              ||       ||||||||||||||||||    ||||    ||      |||   ||||||||||||
Db           600 ACNNTACTGTGGATATGCCTCCCCCCTGACTTANNGCATATNGTT--CATCCCTGTACCT 657

Qy       6190 GAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTT 6249
              | |||||||  |  ||||||||||||||||||||||||||||  ||||||||||||||||||||
Db           658 G-ACAATGGGAACACTTCAGCACAGAAATAAACACCACTTTCGTTTTAGTAGGACCTCTT 716

Qy       6250 GTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACT 6309
              |||||||    ||  || ||||||||||  ||||||  |||||  ||  ||||||||||  ||||
Db           717 GTTTCCATCTGGG--ATTACCCATACCTAAAACCTNACCTGGGT-AAATTTAGCATTACT 773

Qy       6310 ACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 6369
              ||  |  |  ||||||  |||||          |  |||||  |||||  |||  ||||||||||||
Db           774 ACTTCCCCAACCACTTCCCATGCATCAGGTGGGGTACCTCCTTCCAACCAAATAGTCTGC 833

Qy       6370 CTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCT 6429
              |||||||||||||  ||||||||  ||||||  ||||  |  |||  |  ||||||||||  ||||||
Db           834 CTACCCTCAGGAA-ATTTTTTGGCTGTGGGACCTTAACCTTTGGTTGTTTGAAAGGCTCT 892

Qy       6430 TCAGAATCTATGTGCTTCCTCTCATT 6455
              |  ||||  ||||  ||    |  |||  |
Db           893 TAAGAAACTATAGGCCCTCCTTCAAT 918


RESULT 35
DA832031
LOCUS       DA832031                761 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA832031 PLACE1 Homo sapiens cDNA clone PLACE1008285 5', mRNA
            sequence.
ACCESSION   DA832031
VERSION     DA832031.1  GI:82376942
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 761)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED-  16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008285"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             9.6%;  Score 718.6;  DB 1;  Length 761;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches  724;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       5835 TTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTG 5894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 TTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTG 60

Qy       5895 TTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAG 5954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAG 120

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
              |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db           121 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCC 180

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 240

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 300

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db           301 CTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACA 360

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db           361 ATGGAACAACTTCAGCACAGAAATAGACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 420

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
```

```
                ||||:||||||
```

```
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .840
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE3000098"
                     /tissue_type="placenta"
                     /clone_lib="PLACE3"
                     /note="Vector: pME18SFL3"

  Query Match             9.5%;  Score 711.9;  DB 1;  Length 840;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  751;  Conservative   69;  Mismatches   20;  Indels   23;  Gaps    6;

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              | :|||||||||||||:|||||||||||:|:|:|:|||||:|||||||||||||||||
Db             1 GGTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT---- 56

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||:|||||||||||||:::|||||||||
Db            57 TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCT 116

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:||||||||||||||||||:|::|||||||||:|||:|
Db           117 ATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATA 176

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||||||:||||||| |||||||:||||:||||||||||:|||||:|||||:|
Db           177 GGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACT 235

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              ||||||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||||
Db           236 AATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGT 295

Qy       3369 GGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACR 3428
              ||||||||||||||                 ||||||||:||||||||||||:||||||||:
Db           296 GGCTTTCCAGGCCC--------------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACA 341

Qy       3429 GGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTAC 3488
              |||||||||||||:||:|||::||||||  |||||||||:||:||||||:||||| |||||
Db           342 GGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGT-CTTAC 399

Qy       3489 ACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT 3548
              ||||:|||:|||||:|||||||||||::||||::|||||||:|||||||||:|||||||
Db           400 ACAGATCCGAGGGATGAGCTTGCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGT 459

Qy       3549 GGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGA 3608
              ||||||||||||:|:||||||||||:|||||||||||||||||||::||||||||||
Db           460 GGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGA 519

Qy       3609 AGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRG 3668
              |||||||||||||||||||:||||||||||||||||||||||||||:||||::|
```

```
Db        520 AGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGG 579

Qy       3669 CATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGC 3728
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||:||||||
Db        580 CATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTACTTAAATGTCAGGC 639

Qy       3729 TCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNC 3788
              ||||||||| |||:||||||||||||:|||||||||||||||||||||||||||||| :| |
Db        640 TCTATTACTAGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCANCCAC 699

Qy       3789 ATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAAC 3848
              |||||||||||||||||||||||||||| :||: ||||||||||||:||||||||||||
Db        700 ATTTCTTCCAGACAATGAAGAAAAGATTAAACCTAACTGTCAACAAGTAATTTCTCAAAC 759

Qy       3849 CTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGT 3908
              |||||||||||||||||||| |:|| | | |||:|||||||||||    | | ||||||
Db        760 CTATGCCACTCGAGGGGACCCTTTANANGGTCCTTTGACTGATCC--GACCTTAACTTGG 817

Qy       3909 ATACTGATGGAAGTTCCTTTGTA 3931
              ||   ||||||||||||||| |
Db        818 ATCTGGATGGAAGTTCCTTTGNA 840


RESULT 38
BI087886
LOCUS       BI087886                771 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602852690F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4993894 5',
            mRNA sequence.
ACCESSION   BI087886
VERSION     BI087886.1  GI:14506216
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11015  row: d  column: 23
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4993894"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_10"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."

  Query Match             9.5%;  Score 710.9;  DB 1;  Length 771;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  734;  Conservative   18;  Mismatches    4;  Indels    3;  Gaps    3;

Qy       6826 GAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACT 6885
              |||||||||:|||||||||:||||||||||||||||||:||||||:||||||||||||||
Db          1 GAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACT 60
```

```
Qy       6886 GGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 6945
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 GGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 120

Qy       6946 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTT 7005
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        121 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTT 180

Qy       7006 GTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACT 7065
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        181 GTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACT 240

Qy       7066 AAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATC 7125
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        241 AAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATC 300

Qy       7126 AAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCA 7185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCA 360

Qy       7186 GGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT-TG 7244
             ||||||||||||||||||:||||| |||||||||||| |||||||||||||||||| ||
Db        361 GGAAGCAGTTAGAGCGGTCGTCGG-CAACCTCCCCAAAAGCACTTAGGTTTTCCTGTGTG 419

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|||||
Db        420 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAG 479

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        480 CCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAC 539

Qy       7365 CTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTA-AAGAAAT 7423
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        540 CTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTACAAGAAAT 599

Qy       7424 AGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 7483
             ||||||||||:||||||:||||||||||||||||||||||||||||||| || |||
Db        600 AGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATTAAACCA 659

Qy       7484 AGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTG 7543
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTG 719

Qy       7544 TTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             | ||||||||||||||||||||||||||||||:||||:
Db        720 TATTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 758


RESULT 39
DA830020
LOCUS       DA830020                792 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA830020 PLACE1 Homo sapiens cDNA clone PLACE1004263 5', mRNA
            sequence.
ACCESSION   DA830020
VERSION     DA830020.1  GI:82330398
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
    PUBMED  16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004263"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             9.4%;  Score 704.7;  DB 1;  Length 792;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches  711;  Conservative   11;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       5602 TTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGT 5661
             |||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||||||
Db          1 TTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGT 60

Qy       5662 ATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATT 5721
             ||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||||
Db         61 ATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATT 120

Qy       5722 GATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCAT 5781
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db         121 GATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCAT 180

Qy       5782 ATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGG 5841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGG 240

Qy       5842 ACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACT 5901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACT 300

Qy       5902 TACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAA 5961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAA 360

Qy       5962 CATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAA 6021
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         361 CATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAA 420

Qy       6022 GGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTA 6081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTA 480

Qy       6082 TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGG 6141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGG 540

Qy       6142 ATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAAC 6201
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         541 ATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAAC 600

Qy       6202 AACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTG 6261
             ||||||||||||||||||||||||||||||||||||| |||||||||||||||||:||
Db         601 AACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTANGACCTCTTGTTTCCAATCTG 660

Qy       6262 GAAATAACCCATACCTCAAACCTCACCT-GTGTAAAATTTAGCAATACTACATACACAAC 6320
             |||||||||||||||||||| ||||||| ||||||||||||||||||||| ||||||
Db         661 GAAATAACCCATACCTCAAANCTCACCTGGTGTAAAATTTAGCAATACTACATNCACAAC 720
```

```
Qy       6321 CAACTCCCAA 6330
              |||   || |
Db        721 CAAACTCCCA 730


RESULT 40
AU138405
LOCUS       AU138405                721 bp    mRNA    linear   EST 16-NOV-2004
DEFINITION  AU138405 PLACE1 Homo sapiens cDNA clone PLACE1008489 5', mRNA
            sequence.
ACCESSION   AU138405
VERSION     AU138405.2  GI:55780428
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 721)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     On Oct 25, 2000 this sequence version replaced gi:10999926.
            Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .721
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008489"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             9.3%;  Score 697.7;  DB 1;  Length 721;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  704;  Conservative    9;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       6144 ATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAA 6203
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db          2 ATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAA 61

Qy       6204 CTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGA 6263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         62 CTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGA 121

Qy       6264 AATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAA 6323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAA 181

Qy       6324 CTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAAT 6383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAAT 241

Qy       6384 ATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTG 6443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 ATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTG 301

Qy       6444 CTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTA 6503
              ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db         302 CTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTA 361

Qy       6504 TGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGG 6563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 TGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGG 421

Qy       6564 AGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTA 6623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 AGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTA 481

Qy       6624 CAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTT 6683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 CAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTT 541

Qy       6684 GCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTY 6743
              ||||||| |||||||||||||||||||||||||:|||||||||||||||||||||||:
Db         542 GCAAGATNAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTT 601

Qy       6744 GCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGT 6803
              |||||||||||:||| |||||||||||||| |||||||||||||||||||||||||| ||
Db         602 GCTAACCGCTGAAAGANGGGGAACCTGTTTAATTTTAGGGGAAGAATGCTGTTATTANGT 661

Qy       6804 TAATCAATCCGGAATCGTCACT-GAGAAAGTTRAAGAAATTCSAGATCGAATACAACGT 6861
              |||||||||||| |||||||| |||||||||:|||||||||:||||||||||||||||
Db         662 TAATCAATCCGGGATCGTCACTGGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGT 720


RESULT 41
DA830049
LOCUS       DA830049                819 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA830049 PLACE1 Homo sapiens cDNA clone PLACE1004321 5', mRNA
            sequence.
ACCESSION   DA830049
VERSION     DA830049.1  GI:82330584
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /clone="PLACE1004321"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             9.2%;  Score 690.7;  DB 1;  Length 819;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  743;  Conservative    8;  Mismatches   24;  Indels    9;  Gaps    4;

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 60

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 120

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCTCAACTGCT 180

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 240

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 300

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 360

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 420

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 480

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 540

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 600

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||:|||||| ||||||||||||
Db        601 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAACAACCCATACCT 660

Qy       6278 CAAACCTCACC---TGTGTAAAATTTAGCAATACTACATACACAA--CCAACTCCCAATG 6332
             |||||||||||   |||  ||| ||    | ||||||||||||||  | || ||||||||
Db        661 CAAACCTCACCTGGTGTTAAAATTTAGCAANTACTACATACACAAACCAAATTCCCAATG 720

Qy       6333 CATCA--GGTGGGTAACTCCTCCCACACAAA--TAGTCTGCCTACCCTCAGGAATATTTT 6388
             |||||  |  |||||||||||||| |  ||  |||||||||||||||| |   ||  |
Db        721 CATCAAGGTGGGGTAACTCCTCCCAAAANAAATTAGTCTGCCTACCCTCAAGGGAATANT 780

Qy       6389 TTGT 6392
             || |
Db        781 TTTT 784


RESULT 42
CN272394
LOCUS       CN272394                719 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600005410 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN272394
VERSION     CN272394.1  GI:47288808
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 719   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."

  Query Match             9.1%;  Score 684;  DB 1;  Length 719;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  704;  Conservative   15;  Mismatches    0;  Indels   38;  Gaps    1;

Qy       6814 GGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAG 6873
             |||||||||||||||||||||:|||||||||:|||||||||||||||||:||||||:|||
Db          1 GGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAG 60

Qy       6874 CTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTC 6933
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 CTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTC 120

Qy       6934 TTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTC 6993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 TTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTC 180

Qy       6994 CTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
             ||||||                                  ||||||||||||||||
Db        181 CTTGTT----------------------------------AATGGAGCCCAAGATG 202

Qy       7054 CAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGAT 7113
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        203 CAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGAT 262

Qy       7114 GTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGC 7173
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGC 322

Qy       7174 CCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGG 7233
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        323 CCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGG 382

Qy       7234 TTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAA 7293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        383 TTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAA 442

Qy       7294 GAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAAC 7353
             ||||||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        443 GAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAAC 502

Qy       7354 TTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAG 7413
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 TTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAG 562

Qy       7414 GTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGA 7473
              ||||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||
Db        563 GTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGA 622

Qy       7474 TATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTAT 7533
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        623 TATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTAT 682

Qy       7534 GGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAAT 7570
              |||||||||||||||||||||||||||||||||||||
Db        683 GGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAAT 719


RESULT 43
CZ450650
LOCUS       CZ450650                799 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF731h04TF Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_31h04, genomic survey sequence.
ACCESSION   CZ450650
VERSION     CZ450650.1  GI:77926912
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .799
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_31h04"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             9.1%;  Score 682.4;  DB 1;  Length 799;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  704;  Conservative   55;  Mismatches   38;  Indels    6;  Gaps    3;

Qy       3424 CAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTC 3483
              | ||:||||||||||||| :||  ||: |      |||||||||:||:||||||:||  ||
Db          3 CCACGGGGCAAGACTTTCCTT--TATTTTCACAGAAAAAACAGAAATAGCTCTAGGTTTC 60

Qy       3484 CTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGAT 3543
              |||||||||:|||:|||||: ||| |||||||::||||::|||||||:|||||||:||||
Db         61 CTTACACAGGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGAT 120

Qy       3544 GTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTA 3603
              ||||||||||||||||:|:|||||||||: |||||||| ||||||||||::|||||
Db        121 GTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTA 180

Qy       3604 TCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTG 3663
              |||||||||||||| ||||||||| :|||||||||||||||||||||||||||:|||
Db        181 TCTGAAGCAGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTG 240

Qy       3664 AAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRT 3723
              ||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        241 AATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATAT 300

Qy       3724 CAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCA 3783
              ||||||||| | ||||||:| ||||||||||:|||||||||||||||||||||||||||||
Db        301 CAGGCTCTACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCA 360

Qy       3784 GYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCT 3843
              |:| |||||||||||| |||||||||||||| ||:||||||||||||||:|||||   |
Db        361 GCCACATTTCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTT 420

Qy       3844 CAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAA 3903
              |||||||| ||||||| ||||||||||:||||:|||||: ||||||||||:|||| ||||
Db        421 CAAACCTACGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAA 479

Qy       3904 CTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGT 3963
              |||||||||||||||||||||||| |||||||||| |||| ||||:||||||||||||
Db        480 CTTGTATACTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGT 539

Qy       3964 GGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCT 4023
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||:||||
Db        540 TGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCT 599

Qy       4024 RGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATAT 4083
              :|||||||||||||||:|||:|  |||| |||||||   |||||:||||||||:||||||
Db        600 GGCAGAACTAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATAT 656

Qy       4084 ATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGA 4143
              |||||||||||||:|:||||||:|||||||| ||||||||||||::|||||||:|:|||
Db        657 ATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGA 716

Qy       4144 AAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATT 4203
              ||||| ||||||||||||:|||:||||||||||||:|||||||||||||||||||:|||
Db        717 AAGGGGATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATT 776

Qy       4204 ATTAYTGGCWGTACAGAAACCTA 4226
              ||||:|||| ||||||||||||
Db        777 ATTATTGGCGGTACAGAAACCTA 799


RESULT 44
DU836611/c
LOCUS       DU836611                832 bp    DNA     linear   GSS 17-JAN-2006
DEFINITION  MUGQ_CH252P012C12T7_GL269_094 CHORI-252 Vervet Monkey Library
            Cercopithecus aethiops genomic clone CH252-12C12, genomic survey
            sequence.
ACCESSION   DU836611
VERSION     DU836611.1  GI:84044798
KEYWORDS    GSS.
SOURCE      Cercopithecus aethiops (African green monkey)
  ORGANISM  Cercopithecus aethiops
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Cercopithecus.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   Leveque,G., Badhwar,A., Villeneuve,A., Forgetta,V., McKee,K.H.,
            Wasserscheid,J. and Dewar,K.
  TITLE     Genome Quebec/Genome Canada Vervet Monkey Genome Map
  JOURNAL   Unpublished (2005)
COMMENT     Other_GSSs: MUGQ_CH252P012C12Sp6_GL269_046
```

```
            Contact: Dewar K
            McGill University and Genome Quebec Innovation Centre
            McGill University
            740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1A4
            Tel: 514 398 3311 x00089
            Fax: 514 398 1795
            Email: ken.dewar@mcgill.ca
            Plate: 12  row: C  column: 12
            Seq primer: T7 : TAATACGACTCACTATAGGG
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .832
                      /organism="Cercopithecus aethiops"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9534"
                      /clone="CH252-12C12"
                      /sex="male"
                      /cell_type="White blood cell"
                      /dev_stage="Adult"
                      /clone_lib="CHORI-252 Vervet Monkey Library"
                      /note="Vector: pTARBAC2.1; Site_1: EcoR1; Site_2: EcoR1;
                      Constructed by Michael Nefedov in Pieter de Jong's
                      laboratory at BACPAC Resources, Children's Hospital in
                      Oakland Research Institute."

  Query Match             9.1%;  Score 682.2;  DB 1;  Length 832;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  726;  Conservative   60;  Mismatches   45;  Indels    8;  Gaps    5;

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
Db        832 TATATACACTAATTAAGGAAACTCAGAAAGCCAACACCCATTTAGTAAGATGGACACCTG 773

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
Db        772 AAGCAGAATTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGTCCCAGTGTTAAG 715

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
Db        714 CTTGCCAACAGGGCAAGACTTTTCTTTACATGTCACAG-AAAAAACAAGAATAGCTCTAG 656

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
Db        655 GAGTCCTAATACAGGTCTGAGGGACAAACTTGCAACCCATGGCATACCTGAGTAAGGAAA 596

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
Db        595 TTGATGTAGTAGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
Db        535 TAGTATCTGAAGCAGTTAAAATAATACAGGAAAGGGATCTTACTGTGTGGATATCTCATG 476

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
Db        475 ATGTGAAC-GCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTA 417

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTA 3778
Db        416 AGTATCAGGCTCTATTCCCTGAAGGACCAGTGCTGGGACTGCACACTTGTGCAACTCTTA 357

Qy       3779 ACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAA 3838
Db        356 ACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAA 297

Qy       3839 TTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACC 3898
Db        296 TTGCTCAAACCTACGCCACTTGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC 237

Qy       3899 TTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT 3958
Db        236 TT-GGCTTATATACTGATGGAAGTTCCTTTGTCAACAAAGAACTTTGAAAAGCAGGGTAT 178

Qy       3959 GCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCT 4018
Db        177 GCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAATGCT 118

Qy       4019 YAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYA 4078
Db        117 CAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAATT---AGAAGAAAAAAGGGTA 61

Qy       4079 AATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATG 4137
Db          60 AATACATATATAGACTCTAAGTATGCTTACCTAGTCCTTCATGCCCATGCAGCAATATG 2


RESULT 45
AU138097
LOCUS       AU138097                 702 bp    mRNA    linear   EST 02-AUG-2002
DEFINITION  AU138097 PLACE1 Homo sapiens cDNA clone PLACE1007839 5', mRNA
            sequence.
ACCESSION   AU138097
VERSION     AU138097.1  GI:10999618
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES              Location/Qualifiers
     source           1. .702
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="PLACE1007839"
                      /tissue_type="placenta"
                      /clone_lib="PLACE1"
                      /note="Vector: pME18SFL3"

  Query Match             9.1%;  Score 677.9;  DB 1;  Length 702;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  683;  Conservative   10;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
Db          1 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 60

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
Db          61 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 120

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
```

```
Db        121 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 180

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 240

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
              |||||||||||||||:|||:|||||||||||||||||| ||||||||||||||||||||||
Db        241 ATTCTTAGTGCCCCCTATGACCATCTACACTGAACAATATTTATACAGTTATGTCATATC 300

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 360

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 420

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 480

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGC 6752
              |||||||||||||||||||||||||||||:||||||||||||||||||||:|||||||||
Db        481 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGC 540

Qy       6753 TGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATC 6812
              |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATC 600

Qy       6813 CGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAK-AGCAGARG 6871
              |||||||||||||||||||||:|||||||||:|||| ||||||||||  |||| |:|
Db        601 CGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATNGAATACAACGTANAAGCANAGG 660

Qy       6872 AGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRA 6909
              |||||||||||||||||||||||||||||||||||:|
Db        661 AGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAA 698


RESULT 46
BZ611989
LOCUS       BZ611989                795 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHADN27TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-22E6, genomic survey sequence.
ACCESSION   BZ611989
VERSION     BZ611989.1  GI:31520550
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-22E6"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             9.0%;  Score 676.1;  DB 1;  Length 795;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  703;  Conservative   57;  Mismatches   33;  Indels    5;  Gaps    4;

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
              ||:|||||||:|| ||||||||||:||:||| :|||||| |||||||||:||:||||||
Db          1 AGCTTGCCAACGGGACAAGACTTTTCTTTATAAGTCACAG-AAAAAACAGGAATAGCTCT 59

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
              :|||||||||||||:|||:|||||:|||||||||||::||||::||||||:||||||
Db         60 AGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCCATGGCATACCTGAGTAAGGA 119

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
              ||:|||||||||||||| ||||:|: |||||||||:||||| ||| | ||||||||||
Db         120 AACTGATGTAGTGGCAAAGAGTTGGCCACATTGTTTACGGGTATTGGCGACAGTAGCAGT 179

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
              ::|||||||||  |||||||| ||||||||:|||||||||||||||||||||||||||||
Db         180 CTTAGTATCTGAAAAAGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCA 239

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
              |||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||||
Db         240 TGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTACT 299

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
              |||||:|||||||||||||||:| |||||||||  :|||| ||||||||||||||||
Db         300 TAAATATCAGGCTCTATTACTTGAAGGTCCAGTGCTGTGACTGCGCACTTGTGCAACTCT 359

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
              ||||||||:| ||||||||||||||||||||||||||||||:||: ||||||||||||:|
Db         360 TAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAATGTAACTGTCAACAAGT 419

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
              |||| |||||||||||||||| || ||||||||:||||:| |||:||||||||||:||
Db         420 AATTGCTCAAACCTATGCCACTTGAAGGGACCTTTTAGAGGGTCCCTTGACTGATCCCGA 479

Qy       3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
              || ||| ||||||||||||||||||||||||||||||||||||||||| ||||||: ||||
Db         480 CC-TCAGCTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGCAGGGT 538

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
              |||||||||||||||||||||||||:|||||||||| ||||||||||||||||||||||
Db         539 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAAACCCCTCACTCCAGGAACTAGTG 598

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAG--RAAAAAG 4074
              ||:||||:||| |||||| ||||:|||:| |||||||||||||||||||  :||||||
Db         599 CTCAGCTGGCAAAACTAACAGCCCTCACTCGGGCACTAGAATTAGGAGAAGGAAAAAAAG 658

Qy       4075 GGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAAT 4134
              ||:|| |||||||||||||||:|:||||||:|||||||| ||||||||||||::|||||
Db         659 GGTAATTATATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAAT 718

Qy       4135 ATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCAT 4194
              |||:|:|||||||||||||  | ||| |||:||||||||||:|||||||||  |||
Db         719 ATGGAGAGAAAGGGAATTCCCTATTTC-GAGGGAACACCTATCAAACATCAGGAAAGCAT 777

Qy       4195 TAGGARATTATTAYTGGC 4212
              |||||:|||||||:||||
Db         778 TAGGAGATTATTATTGGC 795


RESULT 47
BE734284
```

```
LOCUS       BE734284                714 bp    mRNA    linear   EST 15-SEP-2000
DEFINITION  601565487F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3840572 5',
            mRNA sequence.
ACCESSION   BE734284
VERSION     BE734284.1  GI:10148276
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM530  row: m  column: 21
            High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source          1. .714
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3840572"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_21"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

  Query Match             8.9%;  Score 667.6;  DB 1;  Length 714;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  703;  Conservative   11;  Mismatches    0;  Indels    4;  Gaps    4;

Qy       6303 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 6362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 60

Qy       6363 AGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAA 6422
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         61 AGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCG-TGTTTGAA 119

Qy       6423 TGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACAC 6482
             ||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||
Db        120 TGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACAC 179

Qy       6483 TGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCT 6542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCT 239

Qy       6543 TCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCAC 6602
             ||| |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        240 TCC-TTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCA-TGGCGGTATCAC 297

Qy       6603 AACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGT 6662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 AACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGT 357

Qy       6663 CGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCR 6722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:

Db        358 CGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCA 417

Qy       6723 AAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGG 6782
             ||||||||||||||||||||:||||||||||||:|||||||||||||||||||||||||||
Db        418 AAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGG 477

Qy       6783 GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAAT 6842
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        478 GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 537

Qy       6843 TCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCT 6902
             ||:|||||||||||||||||:||||||:|||||||||||||||||||||||||||||||||
Db        538 TCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCT 597

Qy       6903 CAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCT 6962
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCT 657

Qy       6963 ACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             |||||||||||||||||||||||:| ||||||||||||||||||||||||||||||
Db        658 ACTCCTCTTTGGACCCTGTATCTTTAA-CTCCTTGTTAACTTTGTCTCTTCCAGAATC 714


RESULT 48
DA834377
LOCUS       DA834377                775 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834377 PLACE4 Homo sapiens cDNA clone PLACE4000496 5', mRNA
            sequence.
ACCESSION   DA834377
VERSION     DA834377.1  GI:83039465
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000496"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             8.9%;  Score 665.3;  DB 1;  Length 775;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  749;  Conservative    2;  Mismatches   23;  Indels    7;  Gaps    7;
```

```
Qy       1655 TTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYT 1714
Db          1 TTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCT 60

Qy       1715 CGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAA 1774
Db         61 CGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAA 120

Qy       1775 AAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGG 1834
Db         121 AAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGG 180

Qy       1835 AAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACAC 1894
Db         181 AAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACAC 240

Qy       1895 TTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAA 1954
Db         241 TTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAA 300

Qy       1955 GGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACT 2014
Db         301 GGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACT 360

Qy       2015 AACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATC 2074
Db         361 AACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCTTC 420

Qy       2075 ACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGAC 2134
Db         421 ACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGG-AGGTTGTCTCCTGGAC 479

Qy       2135 ACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTC 2194
Db         480 ACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTC 539

Qy       2195 ACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAG 2254
Db         540 ACTA-TCTGAGGGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAG 596

Qy       2255 TTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCA 2314
Db         597 TTATG-ACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCA 655

Qy       2315 CTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAG 2374
Db         656 CTACCTTGNTAGGGAGAGACATTCTAGC-AAAGCAGGGGCCATTATTCACCTG-ACATNG 713

Qy       2375 GAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGG 2434
Db         714 GNGAAGGAACCCCCCGTTGGTGGCCCCTGCTTTGAGGAAGGAATTAATCCTTGAAGCTTG 773

Qy       2435 G 2435
              |
Db         774 G 774
```

```
RESULT 49
DA834266
LOCUS       DA834266                 850 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834266 PLACE4 Homo sapiens cDNA clone PLACE4000151 5', mRNA
            sequence.
ACCESSION   DA834266
VERSION     DA834266.1  GI:83038598
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 850)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .850
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000151"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"
```

```
  Query Match             8.8%;  Score 657.5;  DB 1;  Length 850;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  799;  Conservative   13;  Mismatches   23;  Indels   17;  Gaps   13;

Qy       1017 CAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGT 1076
Db          1 CAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGT 59

Qy       1077 GCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAA 1136
Db            60 GCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAA 119

Qy       1137 YCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAG 1196
Db           120 CCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAG 179

Qy       1197 AGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAA 1256
Db           180 AGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAA 238

Qy       1257 CTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGAC 1316
Db           239 CTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGAC 296

Qy       1317 AACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTC 1376
Db           297 AACA-AAGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTC 355

Qy       1377 ATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGT 1436
Db           356 ATTGGG---ACACAGAATCAGAACAT-TGAGATTGGTGCTGCAGACATTTGCTAACTTGT 411

Qy       1437 GTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCAC 1496
Db           412 GTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCAC 470

Qy       1497 CATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAT 1556
Db           471 CATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAT 529

Qy       1557 TGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTA 1616
Db           530 TGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTA 589
```

```
Qy       1617 AGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCC 1675
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db        590 AGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCGTAAGCC 649

Qy       1676 CGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATC 1735
             ||||||||||||||||||||||||||||||||||||:|||| ||||||||||||||||||
Db        650 CGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGG-TTTTTATAATAGAGATC 708

Qy       1736 AGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATG 1795
             |||||||||| |||||||||||||||| |||||||||||||||| ||||||||||||||
Db        709 AGGAGGAGCA-GCGGAACAGGACAAACNGGATTAAAAAAAAGGCCCCCGCTTTAGTCATG 767

Qy       1796 ACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCT--GGGCAAATTGAA 1853
             | ||| | |||||||||||||||||||||||||||| || ||| ||  || |||||||||
Db        768 A-CCTNANGCAAGTGGACTTTGGAGGCTCTGGAAAAAGGGAAAACTTGGGCCAAATTGAA 826

Qy       1854 TGCCTAATAGGG 1865
             ||||||| | ||
Db        827 TGCCTAAAAAGG 838
```

RESULT 50
DA830985

| | |
|---|---|
| LOCUS | DA830985 717 bp mRNA linear EST 13-NOV-2005 |
| DEFINITION | DA830985 PLACE1 Homo sapiens cDNA clone PLACE1006153 5', mRNA sequence. |
| ACCESSION | DA830985 |
| VERSION | DA830985.1 GI:82339216 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 717) |
| AUTHORS | Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. |
| TITLE | Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes |
| JOURNAL | Genome Res. 16 (1), 55-65 (2006) |
| PUBMED | 16344560 |
| COMMENT | Contact: Takao Isogai<br>FLJ Project (HRI Team)<br>Helix Research Institute<br>2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan<br>Tel: 81-438-52-3975<br>Fax: 81-438-52-3986<br>Email: flj-cdna@nifty.com<br>HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute (HRI); cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and HRI. |

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .717 |
| | /organism="Homo sapiens" |
| | /mol_type="mRNA" |
| | /db_xref="taxon:9606" |
| | /clone="PLACE1006153" |
| | /tissue_type="placenta" |
| | /clone_lib="PLACE1" |
| | /note="Vector: pME18SFL3" |

Query Match 8.7%; Score 650; DB 1; Length 717;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 680; Conservative 5; Mismatches 25; Indels 2; Gaps 2;

```
Qy       5831 CTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCT-GGAGGACTTGGAGTCACT 5889
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          1 CTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGGAGGACTTGGAGTCACT 60

Qy       5890 GTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAG 5949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAG 120

Qy       5950 GCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCT 6009
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         121 GCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCT 180

Qy       6010 AGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGC 6069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGC 240

Qy       6070 CTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCT 6129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCT 300

Qy       6130 ACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCT 6189
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         301 ACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCT 360

Qy       6190 GAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTT 6249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         361 GAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCCCTT 420

Qy       6250 GTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACT 6309
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACT 480

Qy       6310 ACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 6369
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         481 ACATACACAACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 540

Qy       6370 CTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCT 6429
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         541 CTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAACCTATCGTTGTTTGAATGGCTCT 600

Qy       6430 TCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAA 6489
             |||||||||||||||||||| || |||||||||||||:|||:|      | ||||
Db         601 TCAGAATCTATGTGCTTCCTCCCAATCTTAGTGCCCCCTATGACATCTANACTGGNACAA 660

Qy       6490 GATTTATAC-AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||| |||||||||| ||| |||||||||||| |||     ||| |
Db         661 GATTTATACAAGTTATGTCANATCNAAGCCCCGCAACANAAGGAGTACCAAT 712
```

RESULT 51
DA834382

| | |
|---|---|
| LOCUS | DA834382 698 bp mRNA linear EST 03-DEC-2005 |
| DEFINITION | DA834382 PLACE4 Homo sapiens cDNA clone PLACE4000505 5', mRNA sequence. |
| ACCESSION | DA834382 |
| VERSION | DA834382.1 GI:83039500 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 698) |
| AUTHORS | Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. |
| TITLE | Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes |
| JOURNAL | Genome Res. 16 (1), 55-65 (2006) |
| PUBMED | 16344560 |

```
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000505"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             8.6%;  Score 646.5;  DB 1;  Length 698;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  681;  Conservative    1;  Mismatches    9;  Indels    3;  Gaps    3;

Qy       1480 TACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGA 1539
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 TACTCAATGATGTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGA 59

Qy       1540 GAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCA 1599
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         60 GAGACTAAGGGAGGCATCGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCA 119

Qy       1600 ACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCA 1658
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        120 ACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAACTGCAGACATTAGAAAAAAACTTCA 179

Qy       1659 AAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGT 1718
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 AAGGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGT 239

Qy       1719 TTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGG 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGG 299

Qy       1779 CCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAA 1838
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 CCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAA 359

Qy       1839 GCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTA 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTA 419

Qy       1899 AAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGA 1958
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        420 AAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGA 479

Qy       1959 ATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACC 2018
             |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||
Db        480 ATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACC 539

Qy       2019 AGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCC 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 AGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCC 599

Qy       2079 TCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTG 2138
             |||||||||||||||||||||||||||||||||||| | ||||||| ||||||||||
Db        600 TCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGG-ANGTTGTCTNCTGGACACTG 658

Qy       2139 GTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGG 2172
             |||||| |||||||||||||||||||||||||||
Db        659 GTGCGGNCTTCTTAGTCTTACTCTTCTGTCCCGG 692
```

```
RESULT 52
CJ463212
LOCUS       CJ463212                769 bp    mRNA    linear   EST 06-OCT-2005
DEFINITION  CJ463212 macaque brain cDNA library QnpA Macaca fascicularis cDNA
            clone QnpA-11334 5', mRNA sequence.
ACCESSION   CJ463212
VERSION     CJ463212.1  GI:77257210
KEYWORDS    EST.
SOURCE      Macaca fascicularis (crab-eating macaque)
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirata,M., Terao,K.,
            Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Macaca fascicuraris cDNA database
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Naoki Osada
            Division of Biomedical Resources
            National Institute of Biomedical Innovation
            Saito-Asagi 7-6-8, Ibaraki, Osaka, 567-0085, Japan
            Tel: 81-72-641-9811
            Fax: 81-72-641-9016
            Email: nosada@nibio.go.jp
            URL: http://genebank.nibio.go.jp/gbank/.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Macaca fascicularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:9541"
                     /clone="QnpA-11334"
                     /sex="male"
                     /tissue_type="brain parietal lobe"
                     /dev_stage="adult"
                     /lab_host="TOP10"
                     /clone_lib="macaque brain cDNA library QnpA"
                     /note="Vector: pME18S-FL3 (Acc.No. AB009864); Site_1:
                     DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand
                     cDNA was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     synthesized using specific 5' and 3' primers and amplified
                     by PCR. The PCR product was digested with SfiI and size
                     selection was performed to exclude fragments <1.5kb.The
                     SfiI-digested PCR product was cloned into distinct DraIII
                     sites of pME18S-FL3. XhoI sites just outside the DraIII
                     sites can be used to isolate the cDNA insert. Libraries
                     were constructed by oligo-capping method."

  Query Match             8.6%;  Score 645.2;  DB 1;  Length 769;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  686;  Conservative    9;  Mismatches   57;  Indels    4;  Gaps    1;

Qy       5611 ASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGT 5670
             |:||||:   :||   :||| |||||||||||||||||||||||||||||||||||||||
Db         14 ACTGTTGGCCTACTGGCTTTGACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGT 73

Qy       5671 AGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCA 5730
             |||||||||:|||:|||||||||||||||||||| || ||| ||||||||||||||||||
Db         74 AGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCGGCTTCCTGGAAATATTGATGCCCCA 133

Qy       5731 TCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGC 5790
             || ||||||||| | |:|||||||| ||||||||||||||||||||||||||||||||||
Db        134 TCATATAGGAGTTTATCTAAGGGAAACCCCACCTTCACTGCCCACACCCATATGCCCCGC 193

Qy       5791 AACTGCTATCACTCTGCCACTCTTTGCATGCAT----GCAAATACTCATTATTGGACAGG 5846
             |||| |||| ||||||||||||||||||||||    |||||||||||||||||||||||
Db        194 AACTACTATAACTCTGCCACTCTTTGCATGCATGCAAGCAAATACTCATTATTGGACAGG 253

Qy       5847 AAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTT 5906
             |||| |||||||| || ||||||||||||||||||| ||||||||||||||||||||||
Db        254 GAAAACGATTAATCATAATTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTT 313
```

```
Qy       5907 CACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGT 5966
             |||||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        314 CACCCATACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGTCAAGAGAAAAACATGT 373

Qy       5967 AAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACT 6026
             |||||||||| |||||||||| ||| |||||||| |||| |||||||| ||||||||||
Db        374 AAAAGAAGTACTCTCCCAACTGACCTGGGTACATAGCACGCCTAGCCCCAACAAAGGACT 433

Qy       6027 AGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAA 6086
             |||||||||||||||||||||||||||||| |||||||||||||| ||||| | ||| |||||
Db        434 AGATCTCTCAAAACTACATGAAACCCTCTGTACCCATACTCGTCTGGTGAACCTGTTTAA 493

Qy       6087 TACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATG 6146
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        494 TACCACCCTTACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGCG 553

Qy       6147 CCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTT 6206
             |||||||||| ||||||:|||||| ||||||||||||||||||||||||||||||||||||
Db        554 CCTCCCCCTGCACTTCAGGCCATACGTTTCAATCCCTGTACCTGAACAATGGAACAACTT 613

Qy       6207 CAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAAT 6266
             |||||||||||||||  ||||||||||||||||||||||||||||||||:|||||||
Db        614 CAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAAT 673

Qy       6267 AACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTC 6326
             |||||||||||| |||||||||||||  |||||| |||| ||| | ||||| |||||||
Db        674 AACCCATACCTGAAACCTCACCTGTGGTAAATTTGACAATCCTATAGACACACCCAACTC 733

Qy       6327 CCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 6362
             ||| ||| ||||||||| |||||| |||||| |||
Db        734 CCAGTGCCTCAGGTGGGATACTCCTTCCACACGAAT 769
```

RESULT 53
BZ607553/c
LOCUS BZ607553 833 bp DNA linear GSS 08-JUN-2003
DEFINITION WHADU07TR Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-24A13, genomic survey sequence.
ACCESSION BZ607553
VERSION BZ607553.1 GI:31516114
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 833)
AUTHORS Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .833
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-24A13"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

Query Match 8.6%; Score 644.3; DB 1; Length 833;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 721; Conservative 47; Mismatches 64; Indels 13; Gaps 6;

```
Qy       2571 AAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTA 2630
             | |  || |||  |||||||| |||||||||||||    |  ||||||| || | |
Db        832 AGGACCTAAAAGCCCCAAGGCCTAGTAAAACCATGCTATAGCCCCTGCAATACTCC---- 777

Qy       2631 GTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGA 2690
                 |||   |  | |||||||||||||||||  |||| |||||||||| |||||||||
Db        776 ---AATTTTAGGAGTACAGAAACCCAGTGGACTATGGA-GGTTAGTGCAAAATCTCAGGA 721

Qy       2691 TTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYT 2750
             |||||||| |||||:|||||||:|||||||||||||||||||:|||||||||:||| || ||| ||
Db        720 TTATCAAT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTGTACCCCGCTT 662

Qy       2751 TCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTC 2810
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||:||| || ||| ||
Db        661 TCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGGTGACTTTTTA 602

Qy       2811 TGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCA 2870
             |||||||||||| |||||||||||||||||||||||||||||||||||| |||| ||||||
Db        601 TGCATCCCTGTACTTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCGAACCCA 542

Qy       2871 RCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTT 2930
             :| |||||||||||||||||:||||||| |||||||||||||||||:||||||||||||
Db        541 ACGTCTCAACTCACCTGGACTGTTTTACCTCAAGGGTTCAGGGATAGCCCCCATCTATTT 482

Qy       2931 GGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRG 2990
             |||||||||||||||||||||||||||||:||:|::|||||||||||||||||||||:|
Db        481 GGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTCGG 422

Qy       2991 TAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAA 3050
             ||  |||||||||||||||:||:|||::|||||||||||||||||||||||||||||||
Db        421 TACATGGATGATTTACTTTTAGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCAA 362

Qy       3051 GCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTC 3110
             ||:|||||:|||||||| ||:|||||||||| |: |||||||||:|:|:|||||:|||
Db        361 GCACTCTTAAATTTCCTTGCTACCTGTGGCTAAAAAGTTTCCAAACCAAAGGCTCAGCTC 302

Qy       3111 TGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGT 3170
             |||||||||||||||||||||||||| ||||:|||||||||| ||||||:||||||||||
Db        301 TGCTCACAGCAGGTTAAATACTTAGGTCTAAAATTATCCAAA-GCACCAGGGCCCTCAGT 243

Qy       3171 GAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGR 3230
             ||| ||:::|||||||||| ||||||||||||||||:|||||||||||||||||||||:
Db        242 GAGCAATGTATCCAGCCTAAACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGA 183

Qy       3231 GRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGC 3290
             |::|||||  ||:|||:|||:|||||| |||:||||||| |||||||  |||:|||||||
Db        182 GGGTTCCTCAGCATAACAGGTTTCTGCAGAATATGGATT-CCCAGGTGCGGCGAAATAGC 124

Qy       3291 CAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATG 3350
             ||| :|||||:||||| ||||||||||||||||||||||||||||||||||:|||||||
Db        123 CAGACCATTATATACATTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATG 64

Qy       3351 GAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCA 3410
             ||:|:||||||:::|||  |||||||||||||  ||| | | |||  |||||||:||||
Db         63 GACACCTGAAGCAGAAGCAGCTTTCCAGGCCCTAAAAAAGGCCCT--AACCCAAGCCCCA 6

Qy       3411 GTGTT 3415
             |||||
Db          5 GTGTT 1
```

RESULT 54
BZ601435/c
LOCUS BZ601435 744 bp DNA linear GSS 08-JUN-2003

```
DEFINITION  WHAAS42TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-5H11, genomic survey sequence.
ACCESSION   BZ601435
VERSION     BZ601435.1  GI:31509897
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-5H11"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             8.6%;  Score 642.6;  DB 1;  Length 744;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  675;  Conservative   60;  Mismatches    9;  Indels    6;  Gaps    6;

Qy       2759 ACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCC 2818
             |||||||||||||||||| |||||:|||| ||||||:|||||||||| ||||||||||
Db        744 ACCAGAGGAAGCAGAGTGG-TTACAGTCCT-GACCTTAAGGATGCCTT-TTCTGCATCCC 688

Qy       2819 TGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCA 2878
             |||||||||||||||||||||||||||||||||||| ||| ||||||:  |||||
Db        687 TGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAATGTCTCA 628

Qy       2879 ACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGC 2938
             |||||||||||||:|||||||||||||||||||||||||:|||||||||||||||||||
Db        627 ACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGC 568

Qy       2939 ATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGA 2998
             |||||||||||||||||:||:|::|||||||||||||||||||||||:|||:|||||
Db        567 ATTAGCCCAAGACTTGAGTCAGTTATCATACCTGGACACTCTTGTCCTTCAGTATGTGGA 508

Qy       2999 TGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTT 3058
             |||||||||||:||:|||::|||||||||||||||||||||||||||||:|||||
Db        507 TGATTTACTTTTAGCTGCCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTT 448

Qy       3059 MAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACA 3118
             :||||||||||:|||||||||||||:||||||||||   :|:|||||:||||||||||
Db        447 AAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAAGAGAAGGCTCAGCTCTGCTCACA 388

Qy       3119 GCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYR 3178
             |||||||||||||||||:||||:|||||||||||||||||:||||||||||||||::
Db        387 GCAGGTTAAATACTTAGGACTAAAATTATCCAAAGGCACCAAGGCCCTCAGTGAGGAATG 328

Qy       3179 YATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCT 3238
             :|||||||||||||||||||||||||||||:|||||||||||||||||||||:|::|||||
Db        327 TATCCAGCCTATACTGGCTTATCCTCATCTCAAAACCCTAAAGCAACTAAGAGAGTTCCT 268

Qy       3239 TGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
             ||||:|||:|||:||||||||||:||||||| |||||||:||||:||||||||||:|||
Db        267 TGGCATAACAGGCTTCTGCCGAATATGGATT-CCCAGGTATGGCAAAATAGCCAGGCCAT 209

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db        208 TATATACAGTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGATACCTG 149

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             |||:::|||||||||||||||| |||||||||||||| ||||||||:||||||||||||
Db        148 AAGCAAAAGTGGCTTTCCAGG-CCCTAAAGAAGGCCTT-AACCCAAGTCCCAGTGTTAAG 91

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :||||||||:|||||||||||||||:||:|||::|||||||||||||||||:||:|||||:|
Db         90 CTTGCCAACGGGGCAAGACTTTTCTTTATACATCACAGAAAAAAACAGAAACAGCTCTGG 31

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             |||||||||||||:|||:|||||: ||||
Db         30 GAGTCCTTACACAGGTCCAAGGGACAAGCT 1


RESULT 55
DA834293
LOCUS       DA834293                701 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834293 PLACE4 Homo sapiens cDNA clone PLACE4000228 5', mRNA
            sequence.
ACCESSION   DA834293
VERSION     DA834293.1  GI:83038753
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000228"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             8.4%;  Score 629.6;  DB 1;  Length 701;
  Best Local Similarity   97.0%;  Pred. No. 0;
```

```
  Matches  685;  Conservative    4;  Mismatches   11;  Indels    6;  Gaps    5;

Qy       1449 CTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGG 1508
             ||||||||||||:|||| ||:||||:||||||||||||||||||||||||||||||||| ||
Db             1 CTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACCATAACACA-GG  58

Qy       1509 GAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTG 1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            59 GAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTG 118

Qy       1569 CCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTC 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            119 CCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTC 178

Qy       1629 AGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACT 1687
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db            179 AGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACT 238

Qy       1688 TAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGG 1747
             ||||||||||||||||||||||||:|||||||||||||| ||||||||||||||||||||
Db            239 TAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAAAGATCAGGAGGAGCAGG 298

Qy       1748 CGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAA 1807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            299 CGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAA 358

Qy       1808 GTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCT 1867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            359 GTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCT 418

Qy       1868 TGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCG 1927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            419 TGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCG 478

Qy       1928 CCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGA 1987
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            479 CCCC-TCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGA 537

Qy       1988 CAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGC 2047
             |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db            538 CAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGC 597

Qy       2048 CTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATT 2107
             |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db            598 CTGGGGCAAGCGCCATCCCATGCCATCACCCTNACAGAGCCCTANGTATGCTTGACCATT 657

Qy       2108 GAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAG 2153
             |||||||| |||| |  || ||||||||||||||| ||||||||
Db            658 GAGGGCCAAGAAG--TTGCTNCTGGACACTGGTGCGGGCTTCTTAG 701
```

```
RESULT 56
DA834387
LOCUS       DA834387                791 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834387 PLACE4 Homo sapiens cDNA clone PLACE4000518 5', mRNA
            sequence.
ACCESSION   DA834387
VERSION     DA834387.1  GI:83039515
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .791
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000518"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"
```

```
  Query Match             8.4%;  Score 629.6;  DB 1;  Length 791;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  736;  Conservative   19;  Mismatches   36;  Indels   16;  Gaps    9;

Qy       2293 CTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGG 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 CTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGG  60

Qy       2353 GCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGA 2412
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db            61 GCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGA 119

Qy       2413 AGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCC 2472
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db           120 AGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCC 178

Qy       2473 GTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCC 2532
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db           179 GTCCTGTTCAAGTTAAACTAAAGGATTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCC 238

Qy       2533 TCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCT 2592
             |||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db           239 TCAGACCCAAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCC 298

Qy       2593 TAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAA 2652
             |||||||||||||||  | |||||||||| | |         |||   |  | |||||||
Db           299 TAGTAAAACCATGCAGTAACCCCTGCAGTACTCC-------AATTTTAGGAGTACAGAAA 351

Qy       2653 CCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCC 2712
             ||||   |||||||| |||||||||||||||||||||||||||||| |||||  |||||  |
Db           352 CCCAACAGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGCTGTTGTTC 409

Qy       2713 TTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAG 2772
             | |||| |||||||||||||||||||||||| ||::|||||||||||||||||||||||
Db           410 CTCTATAGCCAGCTGTACCTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAG 469

Qy       2773 AGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACT 2832
             ||||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||||
Db           470 AGTGGTTTACAGTCCTGGACCTTCAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACT 529

Qy       2833 CTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTR 2892
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||:
Db           530 CTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCAACATCTCAACTCACCTGGACTA 589

Qy       2893 TTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACT 2952
             |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db           590 TTTTACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTGGCCAGGCATTAGCCCAAGACT 649

Qy       2953 TGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRG 3012
```

```
                ||||:||:|::|||||||||||||  ||||||||||:||| ||||||||||||||||:|
Db        650 TGAGCCAATCCTCATACCTGGACA--CTTGTCCTTCGGTAAGTGGATGATTTACTTTTGG 707

Qy       3013 CYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYA 3072
              |:||| :|||||||||||| ||||||||||||||||||:|||||:|||||||| | :
Db        708 CCGCC-ATTCAGAAACCTTG-GCCATCAAGCCACCCAAGCGCTCTTCAATTTCCTNGNT- 764

Qy       3073 CCTGTGGCTACAWGGTTTCCAAACSAR 3099
              |||| ||||||  |||| | || | |:
Db        765 CCTGNGGCTACTGGGTTCCAAACCAAA 791


RESULT 57
CR765248
LOCUS       CR765248                851 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp468N179_r1 468 (synonym: phrt1) Pongo pygmaeus cDNA clone
            DKFZp468N179 5', mRNA sequence.
ACCESSION   CR765248
VERSION     CR765248.1  GI:52605121
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 851)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp468N179) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp468N179
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .851
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp468N179"
                     /tissue_type="heart"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="468 (synonym: phrt1)"
                     /note="Vector: pSport1_SfiI; Site_1: SfilA; Site_2: SfilB"

  Query Match             8.4%;  Score 627.1;  DB 1;  Length 851;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  719;  Conservative   60;  Mismatches   66;  Indels   11;  Gaps    9;

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::|||||||||||||| |||  ||  |
Db          1 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGACCCTAAAGAA 60

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||| :||||||| ||||:||||||||:|||||||||||||:||:|||:
Db         61 GGCCCT--AACCCAAACCCCAGTGGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATAT 118

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| || ||||||:||:||||||:|||||||||||||:|| :||||||:||||||
Db        119 GTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGATGAGCTT 177

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              |||||| :||||::||||| :||||||||:|||||||||||||| | |||:|:| ||||
Db        178 GCAACCAGTGGCATACCTGGGTAAGGAAACTGATGTAGTGGCAAAAGATTGGCCTTATTG 237

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:||||| |||||||||||||||::|||||||||||||||||||| ||||||||
Db        238 TTTATGGGTAATGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT-ATACAGGG 296

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| ||||||||:|||||::|||||||||||||||||| ||
Db        297 AAGAGATCTTACTGTGTGGACGTCTCATGATGTGAATGGCATACTCACTGCTAAAGGCGA 356

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
              |||||||||| |||||||: ||||||||||:|||||||||||||||||||:|||||||
Db        357 CTTGTGGCTGCCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGT 416

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
              |||||:|||| ||||||| ||||||||||||||:|   |||||||||||||||||||
Db        417 GCTGCGACTGCGCACTTGCACAACTCTTAACCCAGCCATGTTTCTTCCAGACAATGAAGA 476

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
              | || |:||:|||| |||||||||:||||| ||||||||||||||||||||||||||||
Db        477 GAGGAAAGAACATAATTGTCAACAAGTAATTGCTCAAACCTATGCCACTCGAGGGGACCT 536

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
              |:||||:||  |:|||||||||| :|||| ||||||||||||||||||||||||||||||
Db        537 TTTAGAGGTGTCCTTGACTGATCATGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 595

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAG-TGGTCAGTGATAATGGAATAYTTGAA 3988
              |||||||||||||| |||||:|||||||||||  ||||||| ||||||||||:|||||
Db        596 TAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGCGGGTCAGTCATAATGGAATACTTGAA 655

Qy       3989 AGTAAT-CCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              |||||| |||||||||||| ||||||||||:|||| ||| |||||||| |:|| :|:|
Db        656 AGTAATCCCCCTCACTCCANGAACTAGTGCTCAGCTCGCAAAACTAATAGNCCTCCCTTG 715

Qy       4048 GGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||||||||||||||| |: ||| ||| |||||| | ||||||||||:|: |||||:
Db        716 GGCACTAGAATTAGGAGAANGGGAAAGGGGNAAATATNTNTACAGACTCTAAGNATGCTT 775

Qy       4107 ACCTAGTCNTCCATGCCCAT--GMRGCAATATGSARAGAAA-GGGAATTCCTAACTTCYG 4163
              |||||| | ||| || || |   ::||| | || |:|| || |||||||||||  ||  |
Db        776 ACCTAGNCCTCCNTGGCCCTTGCCGGCATTTTGAAGAGNAAGGGGAATTCCTANTTTTGG 835

Qy       4164 AGRGAACACCTATCAM 4179
              ||:| ||| || |||:
Db        836 AGGGGACAGCTTTCAA 851


RESULT 58
DA834354
LOCUS       DA834354                724 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834354 PLACE4 Homo sapiens cDNA clone PLACE4000427 5', mRNA
            sequence.
ACCESSION   DA834354
VERSION     DA834354.1  GI:83039320
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
```

```
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000427"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             8.4%;  Score 626.7;  DB 1;  Length 724;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  693;  Conservative   11;  Mismatches   19;  Indels    9;  Gaps    6;

Qy        704 TATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTA 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTA 60

Qy        764 CAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTC 823
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTC 120

Qy        824 TTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGC 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGC 180

Qy        884 CTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCC 943
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db        181 CTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCC 240

Qy        944 CCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAG 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAG 300

Qy       1004 TGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAG 1063
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        301 TGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAG 359

Qy       1064 CCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTA 1123
              |||||||||||||:||||::||||||||||||||||||||||||||||||||||||||||
Db        360 CCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTA 419

Qy       1124 AATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTG 1183
              |||||||||||||:|||||||||||||:||||||||||||||||||||||||||||||||
Db        420 AATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTG 479

Qy       1184 ATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGA 1243
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        480 ATCTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGA 538

Qy       1244 AGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAAT 1303
              |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        539 AGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAAT 598

Qy       1304 GGATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCC 1363
              |    ||||||||||||| |||||||| ||||||||||||||||||||:|||||| ||
Db        599 G--ATAGGATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTNCC 655

Qy       1364 AGTCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACA 1423
              ||||||:||||||||||||   ||| | |     ||||| ||| | ||||||||||||||||
Db        656 AGTCTAGACCCTCATTGGG---ACACAGAATCAGAACAT-GGANAATGGTGCTGCAGACA 711

Qy       1424 TTTGCTAACTTG 1435
              ||||  |  | ||
Db        712 TTTGTNACTTGG 723


RESULT 59
CZ449910/c
LOCUS       CZ449910                757 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF730g17TR Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_30g17, genomic survey sequence.
ACCESSION   CZ449910
VERSION     CZ449910.1  GI:77926157
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_30g17"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             8.4%;  Score 625.6;  DB 1;  Length 757;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  661;  Conservative   48;  Mismatches   37;  Indels    4;  Gaps    4;

Qy       2646 ACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCC 2705
              |||||||||| ||||||||| | |||||||||||||||||||||||| || ||||
Db        756 ACAGAAACCCAATGGACAGTGCA-GGTTAGTGCAAGATCTCAGGATTATCGAT-GAGGTT 699

Qy       2706 GTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAG 2765
              ||||||| || |||||||||||||||| |||||||||| ||::|||||||||||||||||
Db        698 GTTGTCCCTTCATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAG 639

Qy       2766 GAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACAT 2825
              |||||||||||||||||:|||||||||:|||||||||| ||||||||||||||||||||
Db        638 GAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACAT 579

Qy       2826 CCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACC 2885
              |||||||||||||||||||||||||||||||||||| |||  ||||||:||||||||||||
Db        578 CCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACATCTCAACTCACC 519

Qy       2886 TGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCC 2945
```

```
                ||||||:||||| ||||| ||||||||||||:|||||||||||||||||||||||||| |
Db        518 TGGACTGTTTTA-CCCAAAGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGTC 460

Qy       2946 CAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTA 3005
                |||||||||||:||:|::|||||||||||||||||||| :|||:||||||||||||
Db        459 CAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGATTTA 400

Qy       3006 CTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTC 3065
                ||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:|||||
Db        399 ATTTTAGCCACCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTT 340

Qy       3066 CTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTT 3125
                |||||:||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||
Db        339 CTCGCCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTT 280

Qy       3126 AAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAG 3185
                |||||:|||||:||||:|||||||||||||||||:|||||||||||||||||:::||||||
Db        279 AAATATTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAG 220

Qy       3186 CCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTA 3245
                ||||||||||||||||||||:|||||||||||||||||||:| :|||||||||:||
Db        219 CCTATACTGGCTTATCCTCATCTCAAAACCCTAAAGCAACTAAGAGTGTTCCTTGGCATA 160

Qy       3246 AYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATAC 3305
                |:|||:||||||||||:| ||||| |||| ||: || :|||||| ||| :|||||:||||
Db        159 ACAGGTTTCTGCCGAATACGGATT-CCCAAGTACGGTGAAATAGACAGACCATTATATAC 101

Qy       3306 ASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRA 3365
                |:|||||||||||||||||||||||||||||||||||:|||||||||:| ||||||:::|
Db        100 ACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGA 41

Qy       3366 AGTGGCTTTCCAGGCCCCTAAAGAAGGCCT 3395
                |||||||||| ||||   ||  | | |||
Db          40 AGTGGCTTTCCTGGCCTTAAAGAAGGCCCT 11


RESULT 60
BX430050/c
LOCUS       BX430050                658 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX430050 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
            3-PRIME, mRNA sequence.
ACCESSION   BX430050
VERSION     BX430050.2  GI:47003301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30776872.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAB002ZD11_CS00181_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE012YJ24"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             8.3%;  Score 622.1;  DB 1;  Length 658;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  638;  Conservative    5;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       6085 AATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAA-ACCCTACTAACTGTTGGAT 6143
                || |||| || || |||   |||| |||| | ||  |||| ||||||||||||||| ||
Db        658 AAAACCATCCACATTGGATACCATAAGGTTTTGGTACAAATACCCTACTAACTGTTGAAT 599

Qy       6144 ATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAA 6203
                ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        598 ATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAA 539

Qy       6204 CTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGA 6263
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        538 CTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGA 479

Qy       6264 AATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAA 6323
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 AATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAA 419

Qy       6324 CTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAAT 6383
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAAT 359

Qy       6384 ATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTG 6443
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 ATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTG 299

Qy       6444 CTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTA 6503
                ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||| |||
Db        298 CTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTA 239

Qy       6504 TGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGG 6563
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 TGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGG 179

Qy       6564 AGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTA 6623
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTA 119

Qy       6624 CAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTT 6683
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTT 59

Qy       6684 GCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGAC 6741
                |||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db          58 GCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAC 1


RESULT 61
BX374329
LOCUS       BX374329                733 bp    mRNA    linear   EST 23-APR-2004
DEFINITION  BX374329 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI018YH07 5-PRIME, mRNA sequence.
ACCESSION   BX374329
VERSION     BX374329.2  GI:46555522
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
```

```
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30434419.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAM002ZC07_AM127_1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI018YH07"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             8.3%;  Score 620.7;  DB 1;  Length 733;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  705;  Conservative   14;  Mismatches   14;  Indels   11;  Gaps    8;

Qy        814 ACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCC 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCC 60

Qy        874 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTA 933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||::||
Db         61 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTA 120

Qy        934 ATAAGGACCCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTG 180

Qy        994 AACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAA 1053
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 AACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAA 239

Qy       1054 TTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACA 1113
              |||||||||||||||||||||||||:||||::|||||||||||||||||||||||||||
Db        240 TTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACA 299

Qy       1114 GACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGA 1173
              ||||||||||||||||||||||:|||||||||||||:|||||||||||||||||||||||
Db        300 GACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 359

Qy       1174 CAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCC 1233
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        360 CAATTCTTTGATCTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCC 418

Qy       1234 AAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAG 1293
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        419 AAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAG 478

Qy       1294 TCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCA 1353
              ||||||||||    |||||||||||| |||||||| ||||||||||||||||||||||||
Db        479 TCAGGTCAATG--ATAGGATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCA 535

Qy       1354 RGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGT 1413
              :|||||||||||||:|||||||||||   ||| | |     ||||| ||||||||||
Db        536 GGCAGTTCCCAGTCTAGACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGT 591

Qy       1414 GCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCT 1473
              |||||||||||||||||||||||||||||:|||||||||||||||||||:|||| ||:|||
Db        592 GCTGCAGACATTTGCTAACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCT 650

Qy       1474 AYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTT 1533
              |:|||||||||||||||||||||||||||||||  ||||||||||||||| ||||||||||
Db        651 ATGAATTACTCAATGATGTCCACCATAACACA-AGGAAGGGAAGAAAATTCTACTGCCTT 709

Qy       1534 TCTGGAGAGACTAAGGGAGGCATT 1557
              ||||||||||||||| |||||||||
Db        710 TCTGGAGAGACTAAAGGAGGCATT 733


RESULT 62
HSA408776
LOCUS       HSA408776               738 bp    DNA     linear   GSS 02-AUG-2001
DEFINITION  Homo sapiens GSS, clone BAC RPCI-11 447F21 Sp6, genomic survey
            sequence.
ACCESSION   AJ408776
VERSION     AJ408776.1  GI:15131033
KEYWORDS    GSS; genomic survey sequence.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Beyer,K.S., Klauck,S.M., Wiemann,S. and Poustka,A.
  TITLE     Construction of a physical map of an autism susceptibility region
            in 7q32.3-q33
  JOURNAL   Gene 272 (1-2), 85-91 (2001)
   PUBMED   11470513
REFERENCE   2  (bases 1 to 738)
  AUTHORS   Beyer,K.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-2001) Beyer K.S., Molecular Genome Analysis,
            Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280,
            Heidelberg, 69120, GERMANY
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q32.3-33"
                     /clone="BAC RPCI-11 447F21 Sp6"
                     /sex="male"
                     /clone_lib="RPCI-11"

  Query Match             8.3%;  Score 618.2;  DB 1;  Length 738;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  654;  Conservative   28;  Mismatches   52;  Indels    4;  Gaps    2;

Qy       4154 CTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCW 4213
              ||||||||: ||:||||||||||||:||||||||||||||||||||:|||||||:||||:
Db          1 CTAACTTCCAAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCT 60

Qy       4214 GTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAA 4273
              |||||||||||||:|||||||:|||||| ||||||:|||||||||| |||||||||:||
Db         61 GTACAGAAACCTAAAGAGGTGGCAGTCTTTCACTGCTGGGGTCATCAGAAAGGAAAGGAA 120

Qy       4274 AGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGAC 4333
              |||||||||:|||:|||:|||||||  :||||||||||:|||||||||||||  ||||
Db         121 AGGGAAATAGAAGGGAACTGCCAAGTGGATATTGAAGCCAAAAGAGCTGCAAGGTGGGAC 180

Qy       4334 CCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACC 4393
              |||||||||||||||||||  ||  ||| |||||| |||||||||| ||| |||||||
Db         181 CCTCCATTAGAAATGCTTACAGAAGGACCCCTAGTATGGGGTAATCCCCTCCAGGAAACC 240

Qy       4394 AAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTC 4450
              |||||||||||||||||||| ||||| |||| |||||||||||| |||    |||||  |||
Db         241 AAGCCCCAGTACTCAGCAGAAGAAATAGAAGGGGGAACCTCACAAGGACATAGTTTCCTC 300

Qy       4451 CCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGG 4510
              ||||| ||| |  |||||||||||||||| |||||||||||||||||| ||| ||||||
```

```
Db        301 CCCTCAGGATGTCTAGCCACTGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGT 360

Qy       4511 AAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATG 4570
              |||||||||||||||||| || ||||||||||||||||| |||||||||||||:|||
Db        361 GAATTACTTAAAACCCTTCACCAGACCTTTCACTTAGGCATTGATAGCACCCATCAGATG 420

Qy       4571 GCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCC 4630
              |||| ||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        421 GCCATATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCC 480

Qy       4631 TGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAA 4690
              |||||| |||||||:|:||| ||||||||||||:|||||||||||||||||||:|||||
Db        481 TGTGAAATGTGCCAAAGAAACAATCCCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAA 540

Qy       4691 ARAACAGGCCATTACCCTGRARAARACTGGCAACTGA-TTTTACCCACAAGCCCAAACCT 4749
              |:|||||||||||||| |:|:||:|||||||||| | |||||||||| ||||||| ||
Db        541 AGAACAGGCCATTACCCAGGAGAAGACTGGCAACTAAGTTTTACCCACATGCCCAAATCT 600

Qy       4750 CAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTC 4809
              ||| ||||||||||||||||||||||||:|| ||||||| |||||||| :| ||||||
Db        601 CAGAGATTTCAGTATCTACTAGTCTGGGTAGATTCTTTCACTGGTTGGGCGGACGCCTTC 660

Qy       4810 CCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCC 4869
              || |||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db        661 CCTTGTAGGACAGAAAAGTCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCT 720

Qy       4870 AGATTCGGACTTCCCCGA 4887
              ||||| ||||||||||||
Db        721 AGATTTGGACTTCCCCGA 738


RESULT 63
AQ351025/c
LOCUS       AQ351025                776 bp    DNA     linear   GSS 07-MAY-1999
DEFINITION  RPCI11-121J5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-121J5,
            genomic survey sequence.
ACCESSION   AQ351025
VERSION     AQ351025.1  GI:4178360
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 776)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-121J5.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .776
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7546300"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-121J5"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             8.3%;  Score 618;  DB 1;  Length 776;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  645;  Conservative   55;  Mismatches   68;  Indels    2;  Gaps    2;

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
              ||  | | |   ||||||||:|||||||||||||:||||||||:|||||||||||||:||:
Db        776 AAAGAAGGCCGGAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTT 717

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
              |||::|||||||||||| |||:||:| ||||:|||||||||| ||||:|||:||| |: |
Db        716 ATATGTCACAGAAAAAACCAGGAATAACTCTAGGAGTCCTTATACAGGTCCAAGGAACCA 657

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
              |||||||   ::|||::|||| |:|||||||:|||||||||||||||||||:|:||
Db        656 GCTTGCAGNTTGTGGCATACCTTAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTC 597

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
              ||||||||:|||||||||  |||||||||||::|||||||||||||||||| |||||||
Db        596 ATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAGATAATAC 537

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
              ||| :|||||||||  |||||||||||||||||:|||| ::|||||||||||||||||||
Db        536 AGGTAAGAGATCTTATCGTGTGGACATCTCATGATGTGACCGGCATACTCACTGCTAAAG 477

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
              |||||||||||||||||||| |: ||| |||||||:||||||| ||| |||||||:||
Db        476 GAGACTTGTGGCTGTCAGACAGCCATTTGCTTAAATATCAGGCTGTATGACTTGAAGGGT 417

Qy       3746 CAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATG 3805
              |||||||| :|||| ||||||||||||||||||| ||||:  |||||||||||||| ||||
Db        416 CAGTGCTGTGACTGTACACTTGTGCAACTCTTAATCCAGCTGCATTTCTTCCAGATAATG 357

Qy       3806 AAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
              ||| |||||||:  :||| |||||||  ||:||| |   ||||||||||  ||| |||||
Db        356 AAGGAAAGATAGGTCATATCTGTCAAGAAGTAAATGTGCAAACCTATGCTGCTCAAGGGG 297

Qy       3866 ACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCC 3925
              || ||:|| |:|||||:||||||||||:|| | ||| |||||||||||||||||| |||
Db        296 ACTTTCTAAAGGTTCCCTTGACTGATCCCGAAC-TCAGCTTGTATACTGATGGAAGCTCC 238

Qy       3926 TTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTT 3985
              |||||||| |||||||| |||| |:||| ||||||||||||||||||||||||||:||
Db        237 TTTGTAGATAAAGGACTTTGAAAGGTGGGATATGCAGTGGTCAGTGATAATGGAATACTT 178

Qy       3986 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYT 4045
              |||||||||||| ||||||| ||||||| |||:||||:|||||| ||||||||:|||:|
Db        177 GAAAGTAATCCCCACACTCCACGAACTAGCGCTCAGCTGGCAGAATTAATAGCCCTCACT 118

Qy       4046 KGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTC-TRARTATGC 4104
              |||||||||||||||||||||: || ||||:| |||||||||||||| |:|:|||||
Db        117 CGGGCACTAGAATTAGGAGAAGGGAACAGGGTACATATATATACAGACTCTTAAGTATGC 58

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
              |:|||||||| ||||||||||  ::|||||||:|:||| ||||||||||
Db         57 TTACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGATAGGGAATTCC 8


RESULT 64
DA830356
LOCUS       DA830356                696 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA830356 PLACE1 Homo sapiens cDNA clone PLACE1004890 5', mRNA
            sequence.
ACCESSION   DA830356
VERSION     DA830356.1  GI:82333345
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 696)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .696
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004890"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             8.2%;  Score 617.3;  DB 1;  Length 696;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches  650;  Conservative   14;  Mismatches   23;  Indels    3;  Gaps    3;

Qy       4357 ACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGA 4416
              ||  ||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db             1 ACAACCCCTAGTATAGGGTAATCCCCTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGA 60

Qy       4417 AACAGAATGGGGAACCTCACGAGG-CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAG 4475
              ||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db            61 AACAGAATGGGGAACCTCACGAGGACAGTTTTCTCCCCTCGGGACGGCTAGCCACTGAAG 120

Qy       4476 AAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAA 4535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAA 180

Qy       4536 CCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAG 4595
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db           181 CCTTTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAG 240

Qy       4596 GCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATC 4655
              |||||||||||||||||||||:|||:|||||||||||||||:|||||||:|:||||||||
Db           241 GCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATC 300

Qy       4656 CCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAAR 4715
              ||||||||:|||||||||||||||||||:|||||:|||||||||||||||||:|:||:
Db           301 CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAG 360

Qy       4716 ACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTG 4775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 ACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTG 420

Qy       4776 GGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGA 4835
              ||||:|||||||||||||||||| ||||||||||||||||||||||||||| ||||||
Db           421 GGTAGATACTTTCACGGGTTGGGCANAGGCCTTCCCCTGTAGGACAGAAAAGGGCCAAGA 480

Qy       4836 GGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACA 4895
              |||||||||| |||||||||||| |||||||||||||||||||||||||||| ||||| |
Db           481 GGTAATAAAGGNACTAGTTCATGANATAATTCCCAGATTCGGACTTCCCCGAAGCTTANA 540

Qy       4896 GAGTGACAATAGCCCTGCTTTCCAGGCCAC-AGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db           541 GAGTGACAATAGCCCTGCTTTCCAGGCCACAAGTAACCCAGGGAGTATCCCAGGCGTTAG 600

Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCAC-AGTCCTCAGGGAAGGTCGAGAAA 5013
              || |||||||||||||||||||||||||||||| | ||  |||||||| |||| ||| |||||
Db           601 GTTTACGATATCACTTACACTGCGCCTGAANGGCANAAGTCCTCAAGGAAAGTCCAGAAA 660

Qy       5014 ATGAATGAAAYACTCAAAGGACATCTAAAA 5043
              || ||||| |:|||||||||  ||| ||||
Db           661 ATTAATGANACACTCAAAGGGGATCCAAAA 690


RESULT 65
BX409304
LOCUS       BX409304                 792 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX409304 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
            5-PRIME, mRNA sequence.
ACCESSION   BX409304
VERSION     BX409304.2  GI:46956516
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30766967.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAE011ZH12_AE01057_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE012YJ24"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             8.2%;  Score 614.501;  DB 1;  Length 792;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  686;  Conservative   10;  Mismatches   85;  Indels    5;  Gaps    3;

Qy       5609 KTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 5668
              :||:||||:|||:||||   || ||||||||||||||||||||||||||||||||||||
Db             1 TTACTGTTCTTTTACCCCTCTTTACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 60

Qy       5669 GTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 5728
              ||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db            61 GTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 120
```

```
Qy       5729 CATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 5788
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        121 CATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 180

Qy       5789 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 5848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 240

Qy       5849 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 5908
              |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        241 AAATGATTAATCCTAGTTGTCCTGGAACACTTGGAGTCACTGTCTGTTGGACTTACTTCA 300

Qy       5909 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 5968
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        301 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGNAA 360

Qy       5969 AAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 6028
              ||||||  |||  | ||||||  |||:|| ||||||||||||  ||||||||||||||||||||
Db        361 AAGAAGNAATNTNCCAACTAACCCGGNTACATGGCACCTNTAGCCCCTACAAAGGACTAG 420

Qy       6029 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 6088
              |||||||||||||||||||||||||||| ||||||||||||||| ||| |||||  ||||
Db        421 ATCTCTCAAAACTACATGAAACCCTCCGNACCCATACTCGCCTGGNAAGTCTATTNAATA 480

Qy       6089 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 6148
              ||||||||||||  ||||||||||| |||  |||||||||||||||| || ||||||||
Db        481 CCACCCTCACTGNCCTCCATGAGGTTTCGTTCCAAAACCCTACTAACTGCTGTATATGCC 540

Qy       6149 TCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 6208
              ||||||||||||||: |||||||  |||||  ||||||||||||||| || ||||  |||||
Db        541 TCCCCCTGAACTTCAGCCCATATGCCTCAATTCCTGTACCTGAACACTGCAACACCTTCA 600

Qy       6209 GCACAGAAATAAACA-CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATA 6267
              ||||||  ||  |||  |  ||  ||||  ||||||||||||||||||||  ||||:||   ||||
Db        601 GCACAGTAACAAATACCCCCTTCTGTTTTAGTAGGACCTCTTGTTTTCAATCTGTTAATA 660

Qy       6268 ACCCATACCTCAAACCTCACCTGTGTAAAATTT--AGCAATACTACATACACAACCAACT 6325
              | | |  ||||||||  | |||||||||  |||||     |     | | |    ||  |||||
Db        661 TCTCTTTCCTCAAACTTTACCTGTGTACAATTTTAGTCTTATTTTCTTCTCCACTCAACT 720

Qy       6326 CCCAATGCATCAGGTGGGTAAC--TCCTCCCACACAAATAGTCTGCCTACCCTCAGGAAT 6383
              |||||  ||||  |  | ||  |  ||||  ||  |||||  ||  |||  | |  |  | |
Db        721 CCCAAGTCATCCAGCTGCTATCCTTCCTTCCCTACAAAAGGTTTGCTTTCTTTTCGAATC 780

Qy       6384 ATTTTT 6389
              | |||
Db        781 TTCTTT 786


RESULT 66
BX352667
LOCUS       BX352667                725 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX352667 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI018YH07 5-PRIME, mRNA sequence.
ACCESSION   BX352667
VERSION     BX352667.2  GI:46290091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30371681.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAM002ZC06_AM126_1&c=4215.r.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI018YH07"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             8.1%;  Score 610.1;  DB 1;  Length 725;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  698;  Conservative   10;  Mismatches   16;  Indels   11;  Gaps    8;

Qy       1090 CCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTAT 1149
              ||   |||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db          1 CCCACCTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTAT 60

Qy       1150 ATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATGTC 1209
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         61 ATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA-ATGTC 119

Qy       1210 ACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGR 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        120 ACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGA 179

Qy       1270 GTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAAAG 1329
              ||||||||||||||||||||||||||||||||||||     |||||||||||| |||||||
Db        180 GTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-AGGAAAG 236

Qy       1330 ANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACACAG 1389
              | ||||||||||||||||||||||:||||||||||||||:||||||||||||   |||
Db        237 AGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCATTGGG---ACAC 293

Qy       1390 AAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGAC 1449
              | |      ||||| ||||||||||||||||||||||||||||||||||||||:||||||
Db        294 AGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTAGAAGGAC 352

Qy       1450 TAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGG 1509
              ||||||||||:|||| ||:||||:|||||||||||||||||||||||||||||||| |||
Db        353 TAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACCATAACACA-GGG 410

Qy       1510 AAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGC 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGC 470

Qy       1570 CTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCA 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 CTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCA 530

Qy       1630 GTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTT 1688
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        531 GTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTT 590

Qy       1689 AGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGC 1748
              |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        591 AGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGC 650

Qy       1749 GGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAG 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        651 GGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGGCAG 710

Qy       1809 TGGACTTTGGAGGCT 1823
              |||||||||||||||
Db
```

```
Db        711 TGGACTTTGGAGGCT 725


RESULT 67
BZ597795/c
LOCUS       BZ597795               750 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHADK08TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-21B15, genomic survey sequence.
ACCESSION   BZ597795
VERSION     BZ597795.1  GI:31506257
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .750
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-21B15"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             8.0%;  Score 601.8;  DB 1;  Length 750;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  658;  Conservative   49;  Mismatches   36;  Indels    8;  Gaps    6;

Qy       2670 GGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTA 2729
             |||||||||||||||||| |||||| ||||    | |  ||| | | ||||||||||||||
Db        743 GGTTAGTGCAAGATCTCA-GATTATGAATG--AGGCTGTGTTCCTCTATACCCAGCTGTA 687

Qy       2730 CCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTG 2789
             |||| |||||||||| ||::|||||||||||||||||||||||| ||||||||:|||||
Db        686 CCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGCGGTTTACAGTCCTG 627

Qy       2790 GACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 2849
             ||||||:|||||||||| |||||||| ||||||||||||||||||||||||||||||||
Db        626 GACCTTAAGGATGCCTTTTTCTGCAACCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 567

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
             |||||||| |||| |||||:| ||| |||||||||||||:||||||||||||||||||
Db        566 TTTGAAGAGCCTTCGAACCCAACGTCTTAACTCACCTGGACTGTTTTACCCCAAGGGTTC 507

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
             |||||||| ||||| ||||||||| ||| |||||||||||||||||:||:|: ||||||
Db        506 AGGGATAG-CCCCAACTATTTGGCTAGGAATTAGCCCAAGACTTGAGCCAATTTTCATAC 448

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
             ||||||||||||||||||:||| | ||||||||||||||:||:|||::||||||||||

Db        447 CTGGACACTCTTGTCCTTCGGTACGCGGATGATTTACTTTTGGCCGCCCATTCAGAAACC 388

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
             |||||||| ||||||||||||||:|||||:||||||||||:|||||||||||||:||||
Db        387 TTGTGCCACCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGGTT 328

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||||||||| ||||||||||||:|||||:|||||||
Db        327 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGTTTAAATACTTAGGGCTAAAATTATCC 268

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||| ||||:::|||||||||||||| |||||||||||||:|
Db        267 AAAGGCACCAGGGCCCTCAGTGGGGAATGTATCCAGCCTATACT-GCTTATCCTCATCCC 209

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db        208 CAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 149

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| :|||||:|||||:|||||||||||||||||||||||
Db        148 -CCCAGGTACGGCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 90

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::|||||||||||||||||   ||  |
Db         89 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 30

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420
             | |||  ||||||||:|||||||||||||:|
Db         29 GGCCCT--AACCCAAGCCCCAGTGTTAAGCT 1


RESULT 68
BG572445
LOCUS       BG572445               790 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602593490F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4720846 5',
            mRNA sequence.
ACCESSION   BG572445
VERSION     BG572445.1  GI:13580098
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1576  row: c  column: 23
            High quality sequence stop: 624.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4720846"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_79"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); 5' and 3' adaptors were used in cloning
                     as follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3'
                     and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
```

C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

Query Match 8.0%; Score 596.101; DB 1; Length 790;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 700; Conservative 11; Mismatches 25; Indels 11; Gaps 9;

```
Qy       6322 AACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGA 6381
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGA 60

Qy       6382 ATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATG 6441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATG 120

Qy       6442 TGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGT 6501
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        121 TGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGT 180

Qy       6502 TATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCA 6561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCA 240

Qy       6562 GGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTAC 6621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTAC 300

Qy       6622 TACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACC 6681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACC 360

Qy       6682 TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCR-AAATCGAAGAGCTTTAGA 6740
             ||||||||||||||||||||||||||||||||||||||||: ||||||||||||||||||
Db        361 TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCACAAATCGAAGAGCTTTAGA 420

Qy       6741 CTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 6800
             ||:||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 480

Qy       6801 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR-AAGAAATTCSAGATCG--AATACA 6857
             ||||||||||||||||||||||||||||||||||: |||||||||:||||||  |
Db        481 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTACAAGAAATTCGAGATCGCAATACAC 540

Qy       6858 ACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCC 6917
             |||||:||||||:|||||||||||||||| |||||||| |||||||  |  |    |
Db        541 ACGTAGAGCAGAGGAGCTTCGAAACACTGCACCCTGGGG-CTCCTCAGC--AATGGATGC 597

Qy       6918 CTGGATTCTCCCCTTCTTAGGACCTCTAGC-AGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||| |||||||||||||||||||| ||||||| |||||||||||||| |||||
Db        598 CTGGATTCT-CCCTTCTTAGGACCTCTAGCAAGCTATACTATTGCTACTCCTC-TTGGAG 655

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||| ||||  |||||||||||||||||| |||:|| || |
Db        656 CCTGTATCTTTAACCTCCTTGTCAACTGAGTCTCTTCCAGAATCGAAGGTGTAAACCT-C 714

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGA 7063
             |||||    || ||||||||||| ||
Db        715 AAATGCGAGCCCAGATGCAGTCCAGGA 741
```

RESULT 69
BI772113
LOCUS BI772113 723 bp mRNA linear EST 25-SEP-2001
DEFINITION 603059196F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208814 5', mRNA sequence.
ACCESSION BI772113
VERSION BI772113.1 GI:15763691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 723)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM11523 row: o column: 23
High quality sequence stop: 722.
FEATURES Location/Qualifiers
source 1. .723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5208814"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

Query Match 7.9%; Score 594.9; DB 1; Length 723;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 657; Conservative 16; Mismatches 47; Indels 7; Gaps 4;

```
Qy       4537 CTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGG 4596
             |||||||||||||| ||||||||||||:|| ||||||||||||||||||||||||||||
Db          1 CTTTCACTTAGGCATTGATAGCACCCATCAGATAGCCAAATCATTATTTACTGGACCAGG 60

Qy       4597 CCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCC 4656
             |||||||||||||||||||||:|||:|||| |||||||:|||||||:|:||||||||||
Db         61 CCTTTTCAAAACTATCAAGCAGATAGTCAGGTCCTGTGAAGTGTGCCAAAGAAATAATCC 120

Qy       4657 CCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARA 4716
             |||||||:|||||||||||||||||||:||||||:||||||| |||||||  :|:||:|
Db        121 CCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCAATTACCCAAGAGAAGA 180

Qy       4717 CTGGCAACT-GATTTTACCCACAAGCCCAA-ACCTCAGGGATTTCAGTATCTACTAGTCT 4774
             ||||||||| ||||||| ||||| ||| || | | |||||||||||||| ||||||||||
Db        181 CTGGCAACTAGATTTTATCCACATGCCAAACATCGCAGGGATTTCAGTGTCTACTAGTCT 240

Qy       4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
             |||||:|||||||||| |||||||||:|||||||||||||||||||||||||  |||||
Db        241 GGGTAGATACTTTCACTGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGTTCCAAG 300

Qy       4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
             |||||||||||||||||||||||||| |||||||||||||||||||||||| |||||||||
Db        301 AGGTAATAAAGGCACTAGTTCATGAAGTAATTCCCAGATTCGGACTTCCCTGAGGCTTAC 360

Qy       4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
             |||||||||||  | ||||||||| ||||||||||||||||||||||||||||||||||
Db        361 AGAGTGACAATGGTCCTGCTTTCAAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 420

Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
             |||||  |||||||||||||||| |||  |||||||| |||||||||||||| |||||
Db        421 GTATAGAATATCACTTACACTGCACCTAGAGGCCACAATCCTCAGGGAAGGTTGAGAA-- 478
```

```
Qy       5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
              |||||||:|||||| |||||||||| |||| ||||||||||||||||||| |||||
Db        479 --AATGAAACACTCAAACGACATCTAAACAAGCTAACCCAGGAAACCCACCTCGCATGGT 536

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
              |||:|||||| |||||||||| |||||||  |||| ||||||||||||||||||||
Db        537 CTGCTCTGTTGTCTATAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAAGCAGGACTTA 596

Qy       5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA- 5193
              |||||||| |||||||||||| || ||||| |||||||||||||| ||||||||| |
Db        597 GCCCATACAGAATGCTGTATGGACGGTCCTTCCTAACCAATGACCTTCTGCTTGACCAAG 656

Qy       5194 AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAA 5253
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 AGACGGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAA 716

Qy       5254 CATTACA 5260
              |||||||
Db        717 CATTACA 723


RESULT 70
BX430055/c
LOCUS       BX430055               814 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX430055 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YI20
            3-PRIME, mRNA sequence.
ACCESSION   BX430055
VERSION     BX430055.2  GI:47006728
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 814)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30778885.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAB002ZG11_CS00184_1&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .814
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE013YI20"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             7.9%;  Score 590.1;  DB 1;  Length 814;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  719;  Conservative    8;  Mismatches   78;  Indels   15;  Gaps    8;

Qy       5959 AAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTAC 6018
              |||| || || |  || ||| || |||| || |:|||  || ||||||     |||||
Db        814 AAACCTGGAACACTAGAAATTTCTCAACCCAACCGGGAGCAAGGCACCCTATACCCCTTA 755

Qy       6019 AAAGGACTAGATCTCTCAAAACTACA-TGAAACCCTCCGTACCCATACTCGCCTGGTAAG 6077
              |||||  || | ||  |||||  ||| | |||||| | |  ||||    |  |  ||||||
Db        754 AAAGGTATAAAACTGCCAAAAAAACACTAAAACCCCCGGGGCCCAATATGCCTGGGTAAG 695

Qy       6078 CCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTG 6137
              || |||||| |||||||||| |||  ||  ||| ||||||||||||||||||||||||
Db        694 CCAATTTAAAACCACCCTCATGGGG--CCCATAGGGCTCGGCCCAAAACCCTACTAACTG 637

Qy       6138 TTGGATATGCCT-CCCCCTGAACTTCARG--CCATATGTTTCAATCCCTGTACCTGAA-C 6193
              |||||||||||| ||||||||||||||:|  | |    ||  |||||||||||||||| |
Db        636 TTGGATATGCCTCCCCCCTGAACTTCAAGGCCAAATGTTTCAAATCCCTGTACCTGAACC 577

Qy       6194 AATGGAACAACTT---CAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 6250
              |||||| ||||||       || | ||  ||   | ||||||||||||||||||||||||
Db        576 AATGGANCAACTTTCAGCACAGAAAATAAACACCAACTTCCGTTTTAGTAGGACCTCTTG 517

Qy       6251 TTTCCAATSTGGAAAT-AACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACT 6309
              ||||||||:|||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        516 TTTCCAATCTGGAAATAAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACT 457

Qy       6310 ACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 6369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 ACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 397

Qy       6370 CTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCT 6429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 CTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCT 337

Qy       6430 TCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAA 6489
              |||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db        336 TCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAA 277

Qy       6490 GATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTT 6549
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GATTTATACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTT 217

Qy       6550 GTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCT 6609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCT 157

Qy       6610 ACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGAC 6669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGAC 97

Qy       6670 TCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGA 6729
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db         96 TCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGA 37

Qy       6730 AGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCT 6769
              |||||||     |:    | |    |:|||||||||||||
Db         36 AGAGCTTAGACTTC----AACCGCTAAAGAGGGGGAACCT 1


RESULT 71
BE019603
LOCUS       BE019603               586 bp    mRNA    linear   EST 06-JUN-2000
DEFINITION  ba84f03.y1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:2907101 5'
            similar to TR:O95244 O95244 ENVELOPE PROTEIN ;contains Alu
            repetitive element;, mRNA sequence.
ACCESSION   BE019603
VERSION     BE019603.1  GI:8279682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40RP from Gibco
            High quality sequence stop: 519.
FEATURES              Location/Qualifiers
     source           1. .586
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2907101"
                      /tissue_type="choriocarcinoma"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_21"
                      /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                      Site_2: EcoRI; cDNA made by oligo-dT priming.
                      Directionally cloned into EcoRI/XhoI sites using the
                      following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                      for average insert size 1.8kb. Library constructed by
                      Ling Hong in the laboratory of Gerald M. Rubin (University
                      of California, Berkeley) using ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies)."

  Query Match             7.7%;  Score 579.6;  DB 1;  Length 586;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  576;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6283 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 6342
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCACCTGTGTTTAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 60

Qy       6343 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 6402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 120

Qy       6403 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 6462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 180

Qy       6463 CCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 6522
             |||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 240

Qy       6523 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 6582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 300

Qy       6583 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 6642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 360

Qy       6643 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 420

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
             ||||||||||||||||||:||||||||||||||||||:||||||||||:|||||||
Db        421 CTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGG 480

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 6822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 540

Qy       6823 ACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||||||||||:|||||||||:||||||||||||||||:|||||
Db        541 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAG 586
```

```
RESULT 72
CZ459267/c
LOCUS       CZ459267                677 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF747h10TR Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_47h10, genomic survey sequence.
ACCESSION   CZ459267
VERSION     CZ459267.1  GI:77937055
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .677
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="MCF7_47h10"
                      /sex="female"
                      /clone_lib="Human MCF7 breast cancer cell line library
                      (MCF7_1)"
                      /note="Vector: pECBAC1; Site_1: HindIII; This library was
                      constructed from MCF7 breast cancer cell line by Amplicon
                      Express (http://www.genomex.com) using their standard
                      procedure."

  Query Match             7.7%;  Score 577.8;  DB 1;  Length 677;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches  594;  Conservative   42;  Mismatches   34;  Indels    2;  Gaps    2;

Qy       2670 GGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTA 2729
             |||||||||||||||||||||||||| | || ||| |||||| | ||||||||||||||
Db        670 GGTTAGTGCAAGATCTCAGGATTATCAGT-GAAGCCATTGTCCCTCTATACCCAGCTGTA 612

Qy       2730 CCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTG 2789
             |||| |||||| ||| ||::||||||||||||||||||| ||||||||||||||:|||||
Db        611 CCTAACCCTTACACTCTGCTTTCCCAAATACCAGAGGACGCAGAGTGGTTTACAGTCCTG 552

Qy       2790 GACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 2849
             |||| |:|| ||||| | |||||||||||||||||||||||||||||||||||||||||
Db        551 GACCCTAAGCATGCCGTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 492

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
             ||||||||| ||||||||||: |||||||||| |||||||:||||||||||||||||||
Db        491 TTTGAAGATCCTTCAAACCCAATGTCTCAACTCATCTGGACTGTTTTACCCCAAGGGTTC 432

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
             ||| ||||:|||||||||||||||||||||||||||||||||||||:||:|::|| |||
Db        431 AGGAATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCCTAC 372

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
             ||||||||||||||||||:|||  ||||||||||||||||:||: ||::||||||||||
Db        371 CTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTTTTAGCCACCCATTCAGAAACC 312
```

```
Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
             ||||||||||||||||||||||||:|||||:|||||||| ||:|||||||||||||:||||
Db        311 TTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTTGCTACCTGTGGCTACAAGGTT 252

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||||||||||||||||||||:||||:|||||||
Db        251 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 192

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:  |||||||||||||:::|||||||||||||||| ||||||||||:|
Db        191 AAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCC 132

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :| :||||||||:|||:|||:|||||||| ||:|||||||
Db        131 AAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACAGGTTTCTGCCAAATATGGATT 72

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:  ||:||||||||||| :|||||:|||||: |||||||||||||| |||||
Db         71 -CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACGAATTAAGGAAACTCAAAAAGC 13

Qy       3330 CAATACCCATTT 3341
              ||||||||||||
Db         12 CAATACCCATTT 1
```

```
RESULT 73
AQ261133
LOCUS       AQ261133                631 bp    DNA     linear   GSS 24-OCT-1998
DEFINITION  CITBI-E1-2506C15.TF CITBI-E1 Homo sapiens genomic clone 2506C15,
            genomic survey sequence.
ACCESSION   AQ261133
VERSION     AQ261133.1  GI:3787657
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2506C15.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2506C15"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

  Query Match             7.7%;  Score 576.2;  DB 1;  Length 631;
  Best Local Similarity   94.0%;  Pred. No. 0;
```

```
  Matches  593;  Conservative    5;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       5958 AAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTA 6017
             |||||| ||||| |||| |||||||||||| |||:||| |||| ||||| ||||||||||
Db          1 AAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCCTAGCCCCTA 60

Qy       6018 CAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAG 6077
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||
Db         61 CAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTGAG 120

Qy       6078 CCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTG 6137
             |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db        121 CCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCTACTAACTG 180

Qy       6138 TTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATG 6197
             |||||| |||||||||||| |||||||:|||||| |||||||||||||| ||||||||||
Db        181 TTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATCCCTGTTCCTGAACAATG 240

Qy       6198 GAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAA 6257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAA 300

Qy       6258 TSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
             |:|||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db        301 TCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAC 360

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTC 6377
             ||||| |||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        361 AACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTC 420

Qy       6378 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        421 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATC 480

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        481 TATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATA 540

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             || | ||||| || ||||||||| |||||||||||||||||||||||||||||||| |
Db        541 CAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAG 600

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGC 6588
             |||| ||||||||||   ||||||||||||
Db        601 AGCAAGAGTGCTAGGCAGACTAGGTACTGGC 631
```

```
RESULT 74
AU134310
LOCUS       AU134310                690 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  AU134310 OVARC1 Homo sapiens cDNA clone OVARC1001704 5', mRNA
            sequence.
ACCESSION   AU134310
VERSION     AU134310.1  GI:10994849
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
```

FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

```
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="OVARC1001704"
                     /tissue_type="ovary, tumor tissue"
                     /clone_lib="OVARC1"
                     /note="Vector: pME18SFL3"
```

Query Match 7.7%; Score 573.5; DB 1; Length 690;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 610; Conservative 45; Mismatches 35; Indels 11; Gaps 4;

```
Qy       3460 AAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTG 3519
Db          1 AAACAGGGATAAGCTCTAGGAGTCCTTACACATGTCTGAGAGACGAGCTTGCAACCTGTG 60

Qy       3520 GCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTA 3579
Db         61 GCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTA 120

Qy       3580 GTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTT 3639
Db        121 ATGGCAGCAATAGCAGTCTTAATATCTAAAGCAGTTAAAATGATACAGTGAAGAGATCTT 180

Qy       3640 ACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTG 3699
Db        181 ACTGTG-----ATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTG 235

Qy       3700 TCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTG 3759
Db        236 TCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGTGCTGCGACTG 295

Qy       3760 TGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAA 3819
Db        296 TGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAA 355

Qy       3820 YATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTT 3879
Db        356 CATAACTGTCAACAAGTAATTGCTCAAACCTATGCCACTTGAGGGGACCTTTTAGGGGTT 415

Qy       3880 CCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGG 3939
Db        416 CCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGG 474

Qy       3940 ACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCT 3999
Db        475 ACTTCAAAATGCGGGGTATGC----CTCAGTGATAATGGAATACTTGAAAGTAATCCCCT 530

Qy       4000 CACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATT 4059
Db        531 CACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAATT 590

Qy       4060 AGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCA 4119
Db        591 AGGAGAAGGAAAAAGGGTAAATATATAT-CAGACTCTAAGTATGCTTACCTAGTCCTTCA 649

Qy       4120 TGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTT 4160
Db        650 TGCCCACGCNNCAATATGGANAGAAAAGGAATTCCTAACTT 690
```

RESULT 75
DA846799
LOCUS DA846799 571 bp mRNA linear EST 11-NOV-2005
DEFINITION DA846799 PLACE6 Homo sapiens cDNA clone PLACE6015377 5', mRNA sequence.
ACCESSION DA846799
VERSION DA846799.1 GI:82041683
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 571)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

```
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6015377"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"
```

Query Match 7.6%; Score 569.4; DB 1; Length 571;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 567; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6003 CACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCA 6062
Db          1 CACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCA 60

Qy       6063 TACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCA 6122
Db         61 TACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCA 120

Qy       6123 AAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCC 6182
Db        121 AAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCC 180

Qy       6183 TGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGG 6242
Db        181 TGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGG 240

Qy       6243 ACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAG 6302
Db        241 ACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAG 300

Qy       6303 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 6362
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 360

Qy       6363 AGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAA 6422
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAA 420

Qy       6423 TGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACAC 6482
                |||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||
Db        421 TGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACAC 480

Qy       6483 TGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCT 6542
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCT 540

Qy       6543 TCCTTTTGTTATAGGAGCAGGAGTGCTAGGT 6573
                |||||||||||||||||||||||||||||||
Db        541 TCCTTTTGTTATAGGAGCAGGAGTGCTAGGT 571


RESULT 76
AQ111469/c
LOCUS       AQ111469                 674 bp    DNA     linear   GSS 29-AUG-1998
DEFINITION  CIT-HSP-2372O20.TR CIT-HSP Homo sapiens genomic clone 2372O20,
            genomic survey sequence.
ACCESSION   AQ111469
VERSION     AQ111469.1  GI:3488126
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 674)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2372O20.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .674
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2372O20"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             7.6%;  Score 566.1;  DB 1;  Length 674;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  592;  Conservative   49;  Mismatches   33;  Indels   11;  Gaps    3;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
                ||| |||||||||| ||::|||||||||||||||||||||||||||||||||:||||||
Db        674 CTACCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGG 615

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
                |||||:|||||||||| | |||||||||||| |||| |||||||||||||||| |||||||
Db        614 ACCTTAAGGATGCCTTTTCCTGCATCCCTGTGCATCTTGACTCTCAATTCTTATTTGCCT 555

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
                |||||||| |||| |||||||: |||||||||||||||||||:||||||||||||||||||
Db        554 TTGAAGATCCTTCGAACCCAATATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCA 495

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
                |||||||:|||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db        494 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 435

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
                ||||||||||||||||||:|||  |||||||||||||||||:||:|||::|||||||||||
Db        434 AGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCCGCCTGTTCAGAAACCT 375

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
                |||||||||||||||||| |||:|||||:|||||||| ||:|||||      |||||
Db        374 TGTGCCATCAAGCCACCCGAGCGCTCTTAAATTTCCTTGCCACCTGT--------GGTTT 323

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
                ||||||:|:|:|||||:|||||||||||||||||||||||||||:||||:|||||||||
Db        322 CCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCA 263

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
                |||| |||||||||||||||||:::|||||||||||||||||||||||||||||||:|
Db        262 AAGGTACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCG 203

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
                ||||||||| || ||||||:|::||| |||||:|||:|||:|||||| |||:|||||||
Db        202 AAACCCTAACGCGACTAAGAGGGTTCTTTGGCATAACAGGCTTCTGCTGAATATGGATT- 144

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
                |||||||:||||:|||||||||  |:||||:|||||:||||||||||||||||||||||||
Db        143 CCCAGGTATGGCAAAATAGCCAAGCCATTATATACACTAATTAAGGAAACTCAGAAAGCC 84

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
                ||||| ||||||:|||||||||:|:||||||:::||  |||||||| ||||||  ||  |
Db         83 AATACTCATTTAGTAAGATGGACACCTGAAGCAGAATCGGCTTTCTAGGCCCTAAAGAAG 24

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTT 3415
                | |||  |||||||||:| |||||||
Db         23 GTCCT--AACCCAAGCCACAGTGTT 1


RESULT 77
DA851356
LOCUS       DA851356                 572 bp    mRNA    linear   EST 10-NOV-2005
DEFINITION  DA851356 PLACE7 Homo sapiens cDNA clone PLACE7001041 5', mRNA
            sequence.
ACCESSION   DA851356
VERSION     DA851356.1  GI:81457854
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
```

```
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES              Location/Qualifiers
     source           1. .572
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="PLACE7001041"
                      /tissue_type="placenta"
                      /clone_lib="PLACE7"
                      /note="Vector: pME18SFL3"

  Query Match             7.6%;  Score 566;  DB 1;  Length 572;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches  557;  Conservative   15;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4583 TTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGC 4642
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||:|||||
Db          1 TTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGC 60

Qy       4643 CARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCAT 4702
             ||:|:||||||||||||||:|||||||||||||||||:||||||:||||||||||
Db         61 CAGAGAAATAATCCCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCAT 120

Qy       4703 TACCCTGRARAARACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGT 4762
             |||||||:|:||:||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACCCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGT 180

Qy       4763 ATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAG 4822
             ||||||||||||||||:||||||||||||||||||:||||||||||||||||||||||
Db        181 ATCTACTAGTCTGGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAG 240

Qy       4823 AAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTC 4882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTC 300

Qy       4883 CCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTAT 4942
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTAT 360

Qy       4943 CCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 420

Qy       5003 AGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCC 5062
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        421 AGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCC 480

Qy       5063 ACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAA 5122
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAA 540

Qy       5123 AAGCAGGACTTAGCCCATACGAAATGCTGTAT 5154
             |||||||||||||||||||||||||||||||
Db        541 AAGCAGGACTTAGCCCATACGAAATGCTGTAT 572


RESULT 78
DA846543
LOCUS       DA846543                571 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA846543 PLACE6 Homo sapiens cDNA clone PLACE6015035 5', mRNA
            sequence.
ACCESSION   DA846543
VERSION     DA846543.1  GI:81817963
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES              Location/Qualifiers
     source           1. .571
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="PLACE6015035"
                      /tissue_type="placenta"
                      /clone_lib="PLACE6"
                      /note="Vector: pME18SFL3"

  Query Match             7.5%;  Score 563.8;  DB 1;  Length 571;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  559;  Conservative   10;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5573 AATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCA 5632
             ||  ||||||||:|||||||||||||||||||||:||:||||:|||:|||||:|||||
Db          1 AAATCCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCA 60

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db         61 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTC 120

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGG 180

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 240

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 300

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 360

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        361 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCC 420

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 480

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 540

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGAT 6143
                |||||||||||||||||||||||||||||||
Db        541 TCTCGGCCCAAAACCCTACTAACTGTTGGAT 571


RESULT 79
AU121943
LOCUS       AU121943                745 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  AU121943 MAMMA1 Homo sapiens cDNA clone MAMMA1001324 5', mRNA
            sequence.
ACCESSION   AU121943
VERSION     AU121943.1  GI:10937178
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MAMMA1001324"
                     /tissue_type="mammary gland"
                     /clone_lib="MAMMA1"
                     /note="Vector: pME18SFL3"

  Query Match             7.5%;  Score 560.3;  DB 1;  Length 745;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  635;  Conservative   62;  Mismatches   47;  Indels   11;  Gaps    8;

Qy       2842 TGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCC 2901
                ||||||||||||||||| ||||||||||:  |||||| ||||||||||:|||||||||
Db          1 TGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAATTCACCTGGACTGTTTTACCCC 60

Qy       2902 AAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCART 2961
                |||||||||||||| ||:|||||||||||||||||||||||||||||||||||||:||:|
Db         61 AAGGGTTCAGGGACAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGT 120

Qy       2962 YMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTT 3021
                ::||| |||||||||||||||||||:|||  |||||||||||||||:||:|||: ||
Db        121 TCTCACACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCTGCCCTTT 180

Qy       3022 CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCT 3081
                ||||||||||||||||||||||||||||||:|||||:|||||||| ||:||||||||||
Db        181 CAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTTGCCACCTGTGGCT 240

Qy       3082 ACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAA 3141
                || :|||||||||||:|:| |||||:|||||||||||||||| ||||||||||||:||||
Db        241 ACCAGGTTTCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAA 300

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
                :|||||| |||||||||:|||||||||||||||:::| |||||| | ||||||||||||
Db        301 AATTATCCTAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATC 360

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAA 3261
                ||||||:| |||||||||||||||||||:|::|||||| ||:|||:|||:|||||| ||
Db        361 CTCATCCC-AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATAACAGGCTTCTGCCAAA 419

Qy       3262 WATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACT 3321
                :||||||| |||||||:||| :|||||||||||:|||||:|||||:|||||||||||||
Db        420 TATGGATT-CCCAGGTATGGTGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACT 478

Qy       3322 CAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCC 3381
                ||||||| |||||||||||:|||| |||| |:||||||: :|||  ||||||||||||
Db        479 CAGAAAG-CAATACCCATTTAGTAAGGTGGAAACCTGAAGCGGAAGCAGCTTTCCAGGCC 537

Qy       3382 CCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTT 3441
                |  || _| | |||  |||||||||:|||||||||||:||||||||:||||||||||
Db        538 CTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT-- 593

Qy       3442 STTYATAYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGG 3501
                   :|||::|||||| |||||||||: |:||||||:|||||||||||||| :|| :|| |
Db        594 -CGTATATATCACAG-AAAAAACAGGGATAGCTCTAGGAGTCCTTACACATGTCTGAGAG 651

Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT-GGCAAAGGGTTG 3560
                |:|||||||| |||::||||::||||||:|||| |||:||||||| | ||| || || ||
Db        652 ACGAGCTTGC-ACCTGTGGCATACCTGAGTAAGAAAACTGATGTANTGGGCNAAAGGNTG 710

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
                :|:|||||| |||:||||| |||  ||| ||||||
Db        711 GCCTCATTGGTTATGGGTAATGGCAGCAATAGCAG 745


RESULT 80
BZ610514
LOCUS       BZ610514                680 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHAAJ87TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-3O6, genomic survey sequence.
ACCESSION   BZ610514
VERSION     BZ610514.1  GI:31519075
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
```

```
    source          1. .680
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="MCF7_1-3O6"
                    /sex="female"
                    /clone_lib="Human MCF7 breast cancer cell line library
                    (MCF7_1)"
                    /note="Vector: pECBAC1; Site_1: HindIII; This library was
                    constructed from MCF7 breast cancer cell line by Amplicon
                    Express (http://www.genomex.com) using their standard
                    procedure."

  Query Match             7.4%;  Score 557.8;  DB 1;  Length 680;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  600;  Conservative   48;  Mismatches   31;  Indels    6;  Gaps    5;

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
             ||:||||||||:||||||||||||| :||:|||::|||||| |||||||||:||:||||||
Db          1 AGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCT 59

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
             :|||||||||||||||:|||:|||||: ||| ||||||||::||||::|||||:||||||
Db         60 AGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGA 119

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
             ||:|||||||||||||||||||:|:||||||||||: |||||||| ||||||||||||||
Db        120 AACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGT 179

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
             ::|||||||||||||||||||| ||||||||| :|||||||||||||||||||||||||||
Db        180 CTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCA 239

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
             |||:|||||::|||||||||||||||||||||||||||||||||||||:|||||||||||
Db        240 TGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACT 299

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
             |||||:|||||||||| | |||||||:| |||||||||:|||||||||||||||||||||
Db        300 TAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCT 359

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAA-CTGTCAACAAR 3835
             ||||||||:| |||||||||||| |||||||||||||||| ||:|||| | | ||||||:
Db        360 TAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATATAACATAACCGGCCAACAAG 419

Qy       3836 TAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYG 3895
             ||||| |||||||||| ||||||| ||||||||||:||||:|||||: |||||||||||:|
Db        420 TAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCG 479

Qy       3896 ACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGG 3955
             ||| |||||||||||||||||||||||||||| |||| |||||| |||| |||||:||||
Db        480 ACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAG-AAAAGGGCTTCAAAAAGTGGGG 537

Qy       3956 TATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGT 4015
             |||||||| |||||||||||||||||:|||||||||||||||||||||  ||||||||
Db        538 TATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCCAGAACTAGT 597

Qy       4016 GCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
             |||:||||:|||||||||||||:|||:|  |||| |||||||| ||  |:|||| ||
Db        598 GCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAATTAGAAG--GGAAAAGGG 655

Qy       4076 GYAAATATATATACAGACTCTRART 4100
             |:||||||||||||||||||:|:|
Db        656 GTAAATATATATACAGACTCTAAGT 680


RESULT 81
DA851363
LOCUS       DA851363                572 bp    mRNA    linear   EST 10-NOV-2005
DEFINITION  DA851363 PLACE7 Homo sapiens cDNA clone PLACE7001049 5', mRNA
            sequence.
ACCESSION   DA851363
VERSION     DA851363.1  GI:81457932
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .572
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7001049"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             7.4%;  Score 554.5;  DB 1;  Length 572;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches  556;  Conservative   15;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       4583 TTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGC 4642
             ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||:|||||
Db          1 TTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGC 60

Qy       4643 CARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCAT 4702
             ||:|:|||||||||||||:|||||||||||||||||||:||||||:||||||||||
Db         61 CAGAGAAATAATCCCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCAT 120

Qy       4703 TACCCTGRARAARACTGGCAAC-TGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAG 4761
             |||||||:|:||:||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 TACCCTGGAGAAGACTGGCAACGTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAG 180

Qy       4762 TATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACA 4821
             |||||||||||||||||||:||||||||||||||||||:||||||||||||||||||||
Db        181 TATCTACTAGTCTGGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACA 240

Qy       4822 GAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTT 4881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTT 300

Qy       4882 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 4941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 360

Qy       4942 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 5001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 420
```

```
Qy       5002 AAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 5061
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        421 AAGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 480

Qy       5062 CACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 5121
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 540

Qy       5122 AAAGCAGGACTTAGCCCATACGAAATGCTGTA 5153
             ||||||||||||||||||||||||||||||||
Db        541 AAAGCAGGACTTAGCCCATACGAAATGCTGTA 572


RESULT 82
DA836641
LOCUS       DA836641                 564 bp    mRNA    linear   EST 08-NOV-2005
DEFINITION  DA836641 PLACE6 Homo sapiens cDNA clone PLACE6002347 5', mRNA
            sequence.
ACCESSION   DA836641
VERSION     DA836641.1  GI:81213784
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6002347"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.4%;  Score 551.3;  DB 1;  Length 564;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  560;  Conservative    3;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5812 CTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCT 5871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCT 60

Qy       5872 GGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGG 5931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGG 120

Qy       5932 GGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACC 5991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACC 180

Qy       5992 SGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACC 6051
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACC 240

Qy       6052 CTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAG 6111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAG 300

Qy       6112 GTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATAT 6171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        301 GTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATAT 360

Qy       6172 GTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCC 6231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCC 420

Qy       6232 GTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGT 6291
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        421 GTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGT 480

Qy       6292 GTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGT-GGGTAACTCC 6350
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        481 GTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGGTAACTCC 540

Qy       6351 TCCCACACAAATAGTCTGCCTACC 6374
             ||||||||||||||||||||||||
Db        541 TCCCACACAAATAGTCTGCCTACC 564


RESULT 83
DA845799
LOCUS       DA845799                 578 bp    mRNA    linear   EST 10-NOV-2005
DEFINITION  DA845799 PLACE6 Homo sapiens cDNA clone PLACE6014097 5', mRNA
            sequence.
ACCESSION   DA845799
VERSION     DA845799.1  GI:81472497
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .578
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6014097"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.4%;  Score 550.6;  DB 1;  Length 578;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  568;  Conservative    7;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       6283 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 6342
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          1 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAATTCCCAATGCATCAGGTGG 60

Qy       6343 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 6402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACC 120

Qy       6403 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 6462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTG 180

Qy       6463 CCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 6522
             |||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGC 240

Qy       6523 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 6582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGT 300

Qy       6583 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 6642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTA 360

Qy       6643 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 420

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
             ||||||||||||||||||:||||||||||||||||||:|||||||||||||:||||||
Db        421 CTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGG 480

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCC-GGAATCGT 6821
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        481 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGGAATCGT 540

Qy       6822 CACTGAG-AAAGTTRAAGAAATTCSAGATCGAATACAA 6858
             ||||||| ||||||:|||||||||:|||||||||||||
Db        541 CACTGAGAAAAGTTAAAGAAATTCGAGATCGAATACAA 578


RESULT 84
AG154259/c
LOCUS       AG154259                665 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-018D14.T7, genomic survey
            sequence.
ACCESSION   AG154259
VERSION     AG154259.1  GI:16683937
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 665)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .665
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-018D14.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"

  Query Match             7.4%;  Score 550.5;  DB 1;  Length 665;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  570;  Conservative   43;  Mismatches   40;  Indels    5;  Gaps    2;

Qy       3497 RAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGG 3556
             :|||||: |||||| ||||::||||::||||||:||| ||||:|||||||||||| |||
Db        665 AAGGGACCAGCTTGGAACCCATGGCATACCTGAGTAAAGAAATTGATGTAGTGGCAGAGG 606

Qy       3557 GTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTA 3616
             ||||:|:||| |||||||:||||||||| ||||||||||||::||||||||||||||||||
Db        605 GTTGGCCTCACTGTTTATGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTA 546

Qy       3617 AAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCA 3676
              ||||||| |||:|||||||||||||||||||||| |    |:| |||::||| ||| |
Db        545 GAATAATAATGGGAAGAGATCTTACTGTGTGGACATAT----ATGCGAACGGCACACTTA 490

Qy       3677 CTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTAC 3736
             ||||||||||||||||||||||||||||||||:|||||||| ||:|||||||||||||||
Db        489 CTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAGATATCAGGCTCTATTAC 430

Qy       3737 TTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTC 3796
             || ||:|||||||||||:| ||||||||||||||||||||||||:| ||||||||||||||
Db        429 TTAAAGGGCCAGTGCTGCAAATGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTC 370

Qy       3797 CAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCA 3856
             ||| |||||||||||||||:||:|||||||| |||||:||||| |||||||||| ||||
Db        369 CAGGCAATGAAGAAAAGATAGAACATAACTGTGAACAAGTAATTGCTCAAACCTACGCCA 310

Qy       3857 CTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGAT 3916
             |||||||||||| ||:||||:|||||:|||||||||||:|||| ||||||||||||||||
Db        309 CTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGAT 251

Qy       3917 GGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAAT 3976
             | |||||||||||||||||||||||||||||||: ||||||||||||||||||||||||
Db        250 GAAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGCAGGGTATGCAGTGGTCAGTGATAAT 191

Qy       3977 GGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATA 4036
             ||||||:|||||||||||||||||||||||||||||||||:||||:|||| ||||||||
Db        190 GGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGGACTAATA 131

Qy       4037 GCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCT 4096
             |||:|||:|  ||||||| |||||||||||:|||||||:|||||||||||||||||||||
Db        130 GCCCTCACTCAGGCACTAAAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCT 71

Qy       4097 RARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
             :|:||||||:|||||||| |||| || || |:: |||||||  :  || |||  ||||
Db         70 AAGTATGCTTACCTAGTCCTCCAAGCGCAGGCAACAATATGAGAGAAAGGGGGTTTCC 13
```

```
RESULT 85
AQ021928
LOCUS       AQ021928                 665 bp    DNA     linear   GSS 09-JUN-1998
DEFINITION  CIT-HSP-2304L9.TR CIT-HSP Homo sapiens genomic clone 2304L9,
            genomic survey sequence.
ACCESSION   AQ021928
VERSION     AQ021928.1  GI:3200664
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 665)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2304L9.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .665
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2304L9"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             7.3%;  Score 548.4;  DB 1;  Length 665;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  575;  Conservative   43;  Mismatches   35;  Indels    4;  Gaps    3;

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :||||||| :|||||||||||||:||:|||::|||||  |||||  ||| ||:|| |||:|
Db         11 CTTGCCAATGGGGCAAGACTTTTCTTTATATGTCACAAAAAAACCCAGTAATAGTTCTAG 70

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||| ||:|| :|||||:||||||||||||::||||::||| ||:|||  |||
Db         71 GAGTCCTTACAAAGGTCTGAGGGACGAGCTTGCAACCTATGGCATACCAGAGTAA-AAAA 129

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
             :|||||||||||||||||||:|:||||||||||:|||||||||  ||||||||||::
Db        130 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCT 189

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
             |||||||||||||| |||||||| |||||||:|||||||||||||||| ||||||||||
Db        190 TAGTATCTGAAGCAATTAAAATGATACAGGGAAGAGATCTTACTGTGTGAACATCTCATG 249

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTT- 3717
             |:| |||::|||||||||||||||||||||||||||||||||||||||| ||:||||||||
Db        250 ATGAGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACGACCGTTTACTTA 309

Qy       3718 -AAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
              ||||:|||| ||||||||||||||:||||||||||||:|||| ||||||||||||||||

Db        310 AAAATATCAGTCTCTATTACTTGAAGGGCCAGTGCTGCGACTGCGCACTTGTGCAACTCT 369

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
             |||||||| | |||||||||||||||||||||||||||||:||:|||||||||||||:|
Db        370 TAACCCAGACACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGT 429

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
             |||| ||||||| |||||||| ||||||||||:|||| |||||:|||||||||||:||
Db        430 AATTGCTCAAACGTATGCCACTTGAGGGGACCTTCTAGAGCTTCCCTTGACTGATCCTGA 489

Qy       3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
             || ||| ||||||||||||||||||||||||||||||||||||||||| |||||||:|||||
Db        490 CC-TCAGCTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGCGGGGT 548

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
             |||||||||||||||||||||||:|||||||||||||||||||||| ||||||| |
Db        549 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAAGAACTAGCG 608

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAA 4073
             ||:||||:||| |||||||||| |||:|  ||||||| |||||||| |||:|||||
Db        609 CTTAGCTGGCAAAACTAATAGCCGTCACTCAGGCACTAAAATTAGGAAAAGGAAAAA 665


RESULT 86
DA861649
LOCUS       DA861649                 567 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA861649 PLACE7 Homo sapiens cDNA clone PLACE7014440 5', mRNA
            sequence.
ACCESSION   DA861649
VERSION     DA861649.1  GI:82131647
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 567)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .567
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7014440"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             7.3%;  Score 548.3;  DB 1;  Length 567;
  Best Local Similarity   97.2%;  Pred. No. 0;
```

```
  Matches  551;  Conservative   14;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       4407 CAGCAGGAGAAACAGAATGGGGAACCTCACGAGG-CAGTTTTCTCCCCTCGGGACGGTTA 4465
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||| ||
Db          1 CAGCAGGAGAAACAGAATGGGGAACCTCACGAGGACAGTTTTCTCCCCTCGGGACGGCTA 60

Qy       4466 GCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACC 4525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACC 120

Qy       4526 CTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTT 4585
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        121 CTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTT 180

Qy       4586 ACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCAR 4645
             ||||||||||||||||||||||||||||||:|||:||||||||||||||:|||||||:
Db        181 ACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAG 240

Qy       4646 ARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTAC 4705
             |:|||||||||||||||:|||||||||||||||||:||||||:||||||||||||||
Db        241 AGAAATAATCCCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTAC 300

Qy       4706 CCTGRARAARACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
             ||||:|:||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 360

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
             |||||||||||||:||||||||||||||||||||:||||||||||||||||||||||||||
Db        361 TACTAGTCTGGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAA 420

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 480

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 540

Qy       4946 AGGCGTTAGGTATACGATATCACTTAC 4972
             |||||||||||||||||||||||||||
Db        541 AGGCGTTAGGTATACGATATCACTTAC 567


RESULT 87
DA849456
LOCUS       DA849456                562 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA849456 PLACE6 Homo sapiens cDNA clone PLACE6018801 5', mRNA
            sequence.
ACCESSION   DA849456
VERSION     DA849456.1  GI:82141039
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6018801"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.3%;  Score 547.5;  DB 1;  Length 562;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  555;  Conservative    6;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5036 ATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCT 5095
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db          1 ATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCT 60

Qy       5096 TAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATG 5155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATG 120

Qy       5156 GAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGA 5215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGA 180

Qy       5216 CATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTG 5275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTG 240

Qy       5276 AGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTC 5335
             ||||||||||||||||||||||||||::||||||||||||||||||||||||||:|||
Db        241 AGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTC 300

Qy       5336 TAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCC 5395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        301 TAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCC 360

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        361 AACTGCGGTTAAAGTGGCTGGAGTAGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 420

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 480

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 540

Qy       5576 CCCCCATGGSCCTCCCTTATCAT 5598
             |||||||:|||||||||||||||
Db        541 -CCCCATGGCCCTCCCTTATCAT 562


RESULT 88
DA845887
LOCUS       DA845887                547 bp    mRNA    linear   EST 10-NOV-2005
DEFINITION  DA845887 PLACE6 Homo sapiens cDNA clone PLACE6014209 5', mRNA
            sequence.
ACCESSION   DA845887
VERSION     DA845887.1  GI:81473746
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6014209"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.3%;  Score 543.8;  DB 1;  Length 547;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  542;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5714 GAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCC 5773
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db          1 GAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCC 60

Qy       5774 ACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTC 5833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTC 120

Qy       5834 ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCT 5893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCT 180

Qy       5894 GTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAA 5953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAA 240

Qy       5954 GAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCC 6013
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        241 GAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCC 300

Qy       6014 CCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGG 6073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGG 360

Qy       6074 TAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTA 6133
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        361 TAAGCCTATTTAATACCACCCTCACTGGGCTCCACGAGGTCTCGGCCCAAAACCCTACTA 420

Qy       6134 ACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAAC 6193
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        421 ACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAAC 480

Qy       6194 AATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTT 6253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTT 540

Qy       6254 CCAATST 6260
             |||||:|
Db         541 CCAATCT 547


RESULT 89
AQ546582/c
LOCUS       AQ546582                647 bp    DNA     linear   GSS 28-MAY-1999
DEFINITION  CITBI-E1-2608G14.TF CITBI-E1 Homo sapiens genomic clone 2608G14,
            genomic survey sequence.
ACCESSION   AQ546582
VERSION     AQ546582.1  GI:4905825
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2608G14.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2608G14"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

  Query Match             7.2%;  Score 542.1;  DB 1;  Length 647;
  Best Local Similarity   87.8%;  Pred. No. 0;
  Matches  570;  Conservative   40;  Mismatches   36;  Indels    3;  Gaps    3;

Qy       3505 AGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYT 3564
               ||||||||||::||||::||||||: ||||||| |||||| |||||||||||||||:|:|
Db        647 AGCTTGCAACCCGTGGCATACCTGAGAAAGGAAAGTGATGTGGTGGCAAAGGGTTGGCCT 588

Qy       3565 CATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATA 3624
             |||||||||: ||||||||  ||||||||||::|||||||||||||||||||||||||||
Db        587 CATTGTTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATA 528

Qy       3625 CAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGT-GAAYRGCATACTCACTGCTAA 3683
             |||||:||||||||||||||||||||||||||:||  ||:  |||||||||| |||
Db        527 CAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTAAAATCACATACTCACTGATAA 468

Qy       3684 AGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARG 3743
             ||||||||||||||||||||||||: ||| |||||||:|||||||||||||||||||:|
Db        467 AGGAGACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGG 408
```

```
Qy       3744 GCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAA 3803
             |||||||||| :|||| ||||||||||||||||||||||||:| ||||||||||||||||
Db        407 GCCAGTGCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAA 348

Qy       3804 TGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGG 3863
             ||||||| |||||:||:|||||||||||||| :| ||| ||||||||||||||||||||||
Db        347 TGAAGAAGAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCCACTCGAGG 288

Qy       3864 GGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTT 3923
             ||||| |:||||:||||| ||||||||||:| || |||||||||||| ||||||||||
Db        287 GGACCATCTAGAGGTTCC-TTGACTGATCCTG-CCCTCAACTTGTATATTGATGGAAGTT 230

Qy       3924 CCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAY 3983
             ||| |||||||||||||||| ||||||:||||||||||||||||||||||||||||||:
Db        229 CCTCTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATAC 170

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| |||:||||:|| ||||||||||||:| |
Db        169 TTGAAAGTAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCGGAACTAATAGCCCTTA 110

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             :|:||||||||||||||| ||||:||||||||:||||||||| ||||||||:|:||||
Db        109 TTTGGGCACTAGAATTAGGGGAAGGAAAAAGGGTAAATATATAAACAGACTCTAAGTATG 50

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
              ||:|||||||| |||||||||  ::||| ||||:|:|| || ||||||
Db         49 CTTACCTAGTCCTCCATGCCCACACAGCACTATGGAGAGGAAAGGAATT 1


RESULT 90
DA847869
LOCUS       DA847869                570 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA847869 PLACE6 Homo sapiens cDNA clone PLACE6016787 5', mRNA
            sequence.
ACCESSION   DA847869
VERSION     DA847869.1  GI:82079672
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6016787"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.2%;  Score 539.9;  DB 1;  Length 570;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches  548;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 73

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 133

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        134 CTCTTCTGGTCCATGTTTCTTACGGCTCGAACTGAGCTTTTGCTCACCGTCCACCACTGC 193

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        194 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 253

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        254 CGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 313

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        314 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 373

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        374 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 433

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        434 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 492

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        493 GAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 552

Qy        541 TTGGGAGCTCTGTGAGCA 558
              ||||||||||||||||||
Db        553 TTGGGAGCTCTGTGAGCA 570


RESULT 91
DA239644/c
LOCUS       DA239644                587 bp    mRNA    linear   EST 31-OCT-2005
DEFINITION  DA239644 BRAWH3 Homo sapiens cDNA clone BRAWH3037044 5', mRNA
            sequence.
ACCESSION   DA239644
VERSION     DA239644.1  GI:78457399
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 587)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
```

```
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .587
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRAWH3037044"
                     /tissue_type="brain"
                     /clone_lib="BRAWH3"
                     /note="Vector: pME18SFL3"

  Query Match             7.2%;  Score 539.8;  DB 1;  Length 587;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches  553;  Conservative    6;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6201 CAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATST 6260
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||:|
Db        587 CAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAATCT 528

Qy       6261 GGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAAC 6320
             |||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||||
Db        527 GGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACAAC 468

Qy       6321 CAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGG 6380
             ||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        467 CAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGG 408

Qy       6381 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTAT 6440
             |||||||||||||||||||||||||||||  |||||||||||| ||| ||||||||
Db        407 AATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCTAT 348

Qy       6441 GTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAG 6500
             ||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||||||
Db        347 GTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATACAA 288

Qy       6501 TTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGC 6560
             ||||||| || ||||||||||||||||||||||||||||||||||||||||  ||||
Db        287 TTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGAGC 228

Qy       6561 AGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTA 6620
             |||| |||||||   | ||||||||||||| ||| ||||||||||||||||||||||||||
Db        227 AGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTCTA 168

Qy       6621 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCAC 6680
             |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db        167 CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTCAC 108

Qy       6681 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGA 6740
             |||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||||
Db        107 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTAGA 48

Qy       6741 CTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAG 6787
             ||:|||||| || ||:||||||||||||||||||||||||||||||||
Db         47 CTTGCTAACTGCAGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAG 1


RESULT 92
DA849357
LOCUS       DA849357                 540 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA849357 PLACE6 Homo sapiens cDNA clone PLACE6018680 5', mRNA
            sequence.
ACCESSION   DA849357
VERSION     DA849357.1  GI:82140231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 540)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6018680"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.2%;  Score 537.6;  DB 1;  Length 540;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  534;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5619 TTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCC 5678
             ||:|||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCC 60

Qy       5679 TYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAG 5738
             |:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAG 120

Qy       5739 GAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTA 5798
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTA 180

Qy       5799 TCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAA 5858
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAA 240

Qy       5859 TCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGG 5918
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGG 300

Qy       5919 TATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAAT 5978
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAAT 360
```

```
Qy       5979 CTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAA 6038
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAA 420

Qy       6039 ACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCAC 6098
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCAC 480

Qy       6099 TGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAA 6158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAA 540


RESULT 93
DA835080
LOCUS       DA835080                 541 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA835080 PLACE6 Homo sapiens cDNA clone PLACE6000237 5', mRNA
            sequence.
ACCESSION   DA835080
VERSION     DA835080.1  GI:83059534
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6000237"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.2%;  Score 537.4;  DB 1;  Length 541;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  532;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5609 KTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 5668
              :||:||||:|||:||||||:||||||||||||||||||||||||||||||||||||||||
Db          1 TTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCA 60

Qy       5669 GTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 5728
              ||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCC 120

Qy       5729 CATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 5788
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db         121 CATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCC 180

Qy       5789 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 5848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAA 240

Qy       5849 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 5908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA 300

Qy       5909 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 5968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAA 360

Qy       5969 AAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 6028
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         361 AAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAG 420

Qy       6029 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 6088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 480

Qy       6089 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 6148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 540

Qy       6149 T 6149
              |
Db         541 T 541


RESULT 94
DU641525
LOCUS       DU641525                 647 bp    DNA     linear   GSS 27-OCT-2005
DEFINITION  Ciuffi-HIV-293T-siLL-37G12.fwdQ=52 Human Integration Site
            Library-Ciuffi-HIV-293T-siLL Homo sapiens genomic, genomic survey
            sequence.
ACCESSION   DU641525
VERSION     DU641525.1  GI:78203895
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Ciuffi,A., Llano,M., Poeschla,E., Hoffmann,C., Leipzig,J.,
            Shinn,P., Ecker,J.R. and Bushman,F.D.
  TITLE     A role for LEDGF/p75 in targeting HIV DNA integration
  JOURNAL   Nat. Med. (2005) In press
COMMENT     Contact: Bushman FD
            Department of Microbiology
            University of Pennsylvania School of Medicine
            402C Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA
            19104-6076, USA
            Tel: 215 573 8732
            Fax: 215 573 4856
            Email: bushman@mail.med.upenn.edu
            The hg17 (May 2004) freeze of the human genome was used.
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /cell_type="293TsiLL (293T stably transfected with 2
                     shRNAs targeting LEDGF/p75)"
                     /clone_lib="Human Integration Site
                     Library-Ciuffi-HIV-293T-siLL"
                     /note="Sequences cloned using TOPO vectors."
```

```
  Query Match             7.2%;  Score 536.8;  DB 1;  Length 647;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  562;  Conservative   37;  Mismatches   46;  Indels    2;  Gaps    2;

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
             |||||||:||:|||||||| |||| :| ||| ||||||||||||||||||||||||||||
Db             2 AAAGATAGAACATAACTGTTAACAGATGATTGCTCAAACCTATGCCACTCGAGGGGACCT 61

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
             |:||||:||| |:||||||||||:|||| ||||||||||||||||||||||||||||||
Db            62 TCTAGAGGTTTCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 120

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
             |||||||||||||||| |||||:|||||||||||||||||||||||||| ||:|||||||
Db           121 TAGAAAAAGGACTTCTAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAGTACTTGAAA 180

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
             |||||||||||||| |||||||||| |||:||||:|||||||||||||||:|||:|  ||
Db           181 GTAATCCCCTCACTTCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGG 240

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
             |||||||||||||||| :||||||||:|||||||||||||||||:|:||||||:|||
Db           241 CACTAGAATTAGGAGAAAGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACC 300

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
             ||||| |||||||||| ::|||||||:|:||| | ||||||||||||||:|||:|||
Db           301 TAGTCCTCCATGCCCACACAGCAATATGGAGAGAGAAGGAATTCCTAACTTCCGAGGGAA 360

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229
             |||||||||:|||||||||||||||||||:|||||||:||||:|||||||||||||:||
Db           361 CACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAG 420

Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289
             ||||||:|||||||||||| : |||||||||| |||||||||:|||||||||||:|||:||
Db           421 AGGTGGCAGTCTTACACTGACAGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGA 480

Qy       4290 AYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGC 4349
             |:||||||||| :|||||||||:|||||||| |||| ||||||||||||||||||||||||
Db           481 ACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAATGCAGGACCCTCCATTAGAAATGC 540

Qy       4350 TTATTAAACTTCCCTTAGTAT-AGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCA 4408
             ||||  ||   | | ||||||  |||||||||| || |||||||| ||| ||||||||||
Db           541 TTATAGAAGGACACCTAGTATGGGGGTAATCCCCTCGGGGAAACCCAGCTCCAGTACTCA 600

Qy       4409 GCAGGAGAAACAGAATGGGGAACCTCACGAGGCAGTTTTCTCCCCTC 4455
             |||||||||| |||||||||||||||||| |||   |  | | | | |
Db           601 GCAGGAGAAATAGAATGGGGAACCTCACCAGGACATAGTTTTCTCCC 647


RESULT 95
DA834637
LOCUS       DA834637                563 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA834637 PLACE5 Homo sapiens cDNA clone PLACE5000234 5', mRNA
            sequence.
ACCESSION   DA834637
VERSION     DA834637.1  GI:83041781
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE5000234"
                     /tissue_type="placenta"
                     /clone_lib="PLACE5"
                     /note="Vector: pME18SFL3"

  Query Match             7.2%;  Score 535.7;  DB 1;  Length 563;
  Best Local Similarity   92.9%;  Pred. No. 0;
  Matches  524;  Conservative   37;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       3653 CTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTT 3712
             |||||||:||||::|||||||||||||||||||||||||||||||||||||||||:|||
Db             1 CTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTT 60

Qy       3713 TACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAA 3772
             |||||||||:|||||||||||||||||:||||||||||:|||||||||||||||||||
Db            61 TACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAA 120

Qy       3773 CTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAAGATARAAYATAACTGTCAAC 3832
             ||||||||||:| |||||||||||||||||||||||||||||:||:||||||||||||
Db           121 CTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAAGATAAAACATAACTGTCAAC 180

Qy       3833 AARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATC 3892
             ||:|||||||||||||||||||||||||||||||||||||||:||||:|||||:||||||
Db           181 AAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGACTGATC 240

Qy       3893 CYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYG 3952
             |:|||| ||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db           241 CCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGTG 299

Qy       3953 GGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACT 4012
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db           300 GGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACT 359

Qy       4013 AGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAA 4072
             ||||||:||||:|||||||||||||||:|||:|:||||||||||||||||||||:||||
Db           360 AGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAA 419

Qy       4073 AGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCA 4132
             ||||:|||||||||||||||||:|:|||||:|||||||| ||||||||||||::|||
Db           420 AGGGCAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCA 479

Qy       4133 ATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCC 4192
             |||||:|:|||||||||||||||||||:|||:||||||||||||:||||||||||||||
Db           480 ATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCC 539

Qy       4193 ATTAGGARATTATTAYTGGCWGTA 4216
             |||||||:|||||||:||||:|||
Db           540 ATTAGGAAATTATTATTGGCTGTA 563


RESULT 96
DA838045
LOCUS       DA838045                538 bp    mRNA    linear   EST 09-NOV-2005
DEFINITION  DA838045 PLACE6 Homo sapiens cDNA clone PLACE6004226 5', mRNA
```

```
            sequence.
ACCESSION   DA838045
VERSION     DA838045.1  GI:81276957
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6004226"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 535.2;  DB 1;  Length 538;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  534;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5688 GTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTS 5747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db          1 GTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTC 60

Qy       5748 TAAGGGAACCCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 5807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGGGAACCCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 120

Qy       5808 CACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTG 5867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTG 180

Qy       5868 TCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGA 5927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGA 240

Qy       5928 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACT 5987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACT 300

Qy       5988 CACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGA 6047
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGA 360

Qy       6048 AACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCA 6107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCA 420

Qy       6108 TGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCC 6167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        421 TGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCC 480

Qy       6168 ATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 6225
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATATGTTTCAGTCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACC 538


RESULT 97
DB062480
LOCUS       DB062480                543 bp    mRNA    linear   EST 10-DEC-2005
DEFINITION  DB062480 TESTI4 Homo sapiens cDNA clone TESTI4002644 5', mRNA
            sequence.
ACCESSION   DB062480
VERSION     DB062480.1  GI:83453890
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI4002644"
                     /tissue_type="testis"
                     /clone_lib="TESTI4"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 534.6;  DB 1;  Length 543;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  528;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        494 ATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGT 553
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGT 60

Qy        554 GAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATGGGAA 613
              |||||||||||||||||:||||||||:||||:::||||||||||:||||||||||||||
Db         61 GAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATGGGAA 120

Qy        614 ACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTG 673
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||:||||||||:|||||||
Db        121 ACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTG 180

Qy        674 ACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCT 733
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCT 240

Qy        734 GAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAA 793
                 ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        241 GAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAA 300

Qy        794 TGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGT 853
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGT 360

Qy        854 AAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATC 913
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATC 420

Qy        914 CCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGAT 973
                 ||||||||||||||::||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGAT 480

Qy        974 AGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCC 1033
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCA 540

Qy       1034 A 1034
                 |
Db        541 A 541


RESULT 98
DB068120
LOCUS       DB068120                539 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB068120 TESTI4 Homo sapiens cDNA clone TESTI4009959 5', mRNA
            sequence.
ACCESSION   DB068120
VERSION     DB068120.1  GI:83098682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI4009959"
                     /tissue_type="testis"
                     /clone_lib="TESTI4"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 534.2;  DB 1;  Length 539;
  Best Local Similarity   97.8%;  Pred. No. 0;
  Matches  527;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        494 ATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGT 553
                 ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGT 60

Qy        554 GAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATGGGAA 613
                 ||||||||||||||||||:||||||||||:||||:::||||||||||:||||||||||||
Db         61 GAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATGGGAA 120

Qy        614 ACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTG 673
                 ||||||||||||||||||||||||||||||||||||||||||:||||||||:|||||||
Db        121 ACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTG 180

Qy        674 ACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCT 733
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCT 240

Qy        734 GAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAA 793
                 ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        241 GAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAA 300

Qy        794 TGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGT 853
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGT 360

Qy        854 AAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATC 913
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATC 420

Qy        914 CCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGAT 973
                 ||||||||||||||::||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGAT 480

Qy        974 AGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCC 1032
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCC 539


RESULT 99
BX357209
LOCUS       BX357209                711 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX357209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 5-PRIME, mRNA sequence.
ACCESSION   BX357209
VERSION     BX357209.2  GI:46307413
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30378136.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI022DE09QP1&c=4215.r.
FEATURES              Location/Qualifiers
     source           1. .711
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DI022YJ18"
                      /tissue_type="PLACENTA COT 25-NORMALIZED"
                      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             7.1%;  Score 533.1;  DB 1;  Length 711;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  537;  Conservative   11;  Mismatches    3;  Indels    1;  Gaps    1;

Qy         57 GTCCCTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAAC 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCCCTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAAC 60

Qy        117 TGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCA 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCA 120

Qy        177 CTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGG 236
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGG 180

Qy        237 TGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGC 296
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        181 TGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGC 240

Qy        297 TTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANT 356
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db        241 TTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGT 300

Qy        357 CACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTAC 416
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        301 CACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTAC 360

Qy        417 CACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGA 476
              |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||
Db        361 CACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGA 419

Qy        477 ATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCAC 536
              ||||||:|||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        420 ATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCAC 479

Qy        537 CATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGAC 596
              ||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||||
Db        480 CATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGAC 539

Qy        597 ATCCMAAGTGAT 608
              ||||:|||||||
Db        540 ATCCAAAGTGAT 551


RESULT 100
DA838367
LOCUS       DA838367                 547 bp    mRNA    linear   EST 09-NOV-2005
DEFINITION  DA838367 PLACE6 Homo sapiens cDNA clone PLACE6004634 5', mRNA
            sequence.
ACCESSION   DA838367
VERSION     DA838367.1  GI:81279419
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES              Location/Qualifiers
     source           1. .547
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="PLACE6004634"
                      /tissue_type="placenta"
                      /clone_lib="PLACE6"
                      /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 532.5;  DB 1;  Length 547;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches  537;  Conservative   10;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5107 TGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTC 5166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTC 60

Qy       5167 ATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCT 5226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCT 120

Qy       5227 TAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAA 5286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAA 180

Qy       5287 AAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCAT 5346
              ||||||||||||||||::||||||||||||||||||||||||||:|||||||||||||||
Db        181 AAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCAT 240

Qy       5347 CCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTA 5406
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        241 CCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTA 300

Qy       5407 AAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAA 5466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAA 360

Qy       5467 AGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGC 5526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGC 420

Qy       5527 GCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSC 5586
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||:|
Db        421 GCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCC 479
```

```
Qy       5587 CTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCC 5646
             ||||||||||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||
Db        480 CTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCC 539

Qy       5647 CCTCCATG 5654
             ||||||||
Db        540 CCTCCATG 547


RESULT 101
DA854810
LOCUS       DA854810                 568 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA854810 PLACE7 Homo sapiens cDNA clone PLACE7005367 5', mRNA
            sequence.
ACCESSION   DA854810
VERSION     DA854810.1  GI:82107218
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7005367"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 531.5;  DB 1;  Length 568;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  524;  Conservative   36;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       3823 AACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCY 3882
             |||||||||||||:||||||||||||||||||||||||||||||||||||:||||:|||||:
Db          1 AACTGTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCT 60

Qy       3883 TTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACT 3942
             ||||||||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACT 119

Qy       3943 TCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCAC 4002
             ||||||||:||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        120 TCGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCAC 179

Qy       4003 TCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGG 4062
             ||||||||||||||||:||||:||||||||||||||||:|||:|:||||||||||||||||
Db        180 TCCAGGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGG 239

Qy       4063 AGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGC 4122
             |||||:|||||||||: |||||||||||||||||||:|:||||||:|||||||| |||||||
Db        240 AGAAGAAAAAAGGGCTAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGC 299

Qy       4123 CCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACA 4182
             |||||::|||||||||:|:||||||||||||||||||||||:|||:|||||||||||||:|||
Db        300 CCATGCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACA 359

Qy       4183 TCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTT 4242
             |||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:||||||
Db        360 TCAGGAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTT 419

Qy       4243 ACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKA 4302
             |||||||:|||||||||| |||||||||:||||||||||:|||:|||:|||||||||:|
Db        420 ACACTGCCGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGAGAACTGCCAAGCAGA 479

Qy       4303 TATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCC 4362
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||||  ||
Db        480 TATTGAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAAAACAACC 539

Qy       4363 CTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             | |||||||||||||||| |||||||||
Db        540 CCTAGTATAGGGTAATCCCCTCCGGGAAA 568


RESULT 102
DA838561
LOCUS       DA838561                 569 bp    mRNA    linear   EST 09-NOV-2005
DEFINITION  DA838561 PLACE6 Homo sapiens cDNA clone PLACE6004864 5', mRNA
            sequence.
ACCESSION   DA838561
VERSION     DA838561.1  GI:81280812
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 569)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                     /clone="PLACE6004864"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 531;  DB 1;  Length 569;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches  549;  Conservative    5;  Mismatches    3;  Indels    2;  Gaps    2;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 71

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 131

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGC-TGAGCTTTTGCTCACCGTCCACCACTG 179
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        132 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCNTGAGCTTTTGCTCACCGTCCACCACTG 191

Qy        180 CTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGT 239
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        192 CTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGT 251

Qy        240 CCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTG 299
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        252 CCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTG 311

Qy        300 CCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCAC 359
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        312 CCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCAC 371

Qy        360 TGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCAC 419
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        372 TGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCAC 431

Qy        420 ATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATA 479
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        432 ATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATA 490

Qy        480 CGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCAT 539
              |||:|||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        491 CGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCAT 550

Qy        540 CTTGGGAGCTCTGTGAGCA 558
              ||||||||||||||||||
Db        551 CTTGGGAGCTCTGTGAGCA 569


RESULT 103
BX952683/c
LOCUS       BX952683                652 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  DKFZp781L23211_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781L23211 5', mRNA sequence.
ACCESSION   BX952683
VERSION     BX952683.1  GI:43430094
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781L23211) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781L23211"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"

  Query Match             7.1%;  Score 530.3;  DB 1;  Length 652;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  566;  Conservative   36;  Mismatches   45;  Indels    3;  Gaps    3;

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||: |||||||||||||:||||| |||||||||||||||| || ||||
Db        652 AGAAAAGATAGAATGTAACTGTCAACAAGTAATTGCTCAAACCTATGCCACTTGAAGGGA 593

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:| |||:|| |||||||||:|||| ||| ||||||||||||||||||||||
Db        592 CCTTTTAGAGGGTCCCTTAACTGATCCCGACC-TCAGCTTGTATACTGATGGAAGTTCCT 534

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| ||||||: ||||||||||||||||||||||||||||:|||
Db        533 TTGTAGAAAAAGGACTTTGAAAAGCAGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 474

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||| ||||||||||||||||||||||:||||:||| |||||| ||||:|||:|
Db        473 AAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCCCTCACTC 414

Qy       4047 GGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCT 4105
              |||||||||||||||||||| |:||||||||:|||||||||||||||||||:|:||||||
Db        413 GGGCACTAGAATTAGGAGAAGGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCT 354

Qy       4106 YACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAG 4165
              :|||||||| ||||||||||::||||||||:|:|||||||||||||||| ||| |||
Db        353 TACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTC-GAG 295

Qy       4166 RGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCT 4225
              :|||||||||||:||||||||||||||||||:|||||||:||||:||||||||||||
Db        294 GGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCT 235

Qy       4226 ARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAA 4285
              :|||||||:|||||||||||||:    |||||| |||||||||:|||||||||||:||
Db        234 GAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGAAATAGAA 175

Qy       4286 GRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAA 4345
              |:|||: ||||||||:|||||||| :|||||||||||||||||||||||||||||||||
Db        174 GGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAA 115

Qy       4346 ATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTAC 4405
              |||||||  ||   ||| |||||| |||||||||  |||||||||||||| ||||||||
Db        114 ATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTCCCAGTAC 55

Qy       4406 TCAGCAGGAGAAACAGAATGGGGAACCTCACGAGGCAGTTTTCTCCCCTC 4455
              ||||||||| ||| ||||| |||||||||||||||    |   |   || | |
Db         54 TCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCC 5


RESULT 104
DA856636
LOCUS       DA856636                559 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA856636 PLACE7 Homo sapiens cDNA clone PLACE7007723 5', mRNA
            sequence.
```

```
ACCESSION   DA856636
VERSION     DA856636.1  GI:82162761
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7007723"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 530;  DB 1;  Length 559;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  548;  Conservative    6;  Mismatches    4;  Indels    2;  Gaps    2;

Qy         14 ATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGTAT 73
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGTAT 60

Qy         74 GGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCA 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCA 120

Qy        134 TGTTTCTTACGGCTCGAGC-TGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCA 192
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTTCTTACGGCTCGAGCNTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCA 180

Qy        193 CCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCT 252
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCT 240

Qy        253 GATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGC 312
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 GATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGC 300

Qy        313 ACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGT 372
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        301 ACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGT 360

Qy        373 TCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATT 432
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        361 TCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATT 420

Qy        433 CCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCAC 492
              |||||||||||||||||||||||:||| ||||||||||||||||||||||:|||||||||
Db        421 CCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCAC 479

Qy        493 CATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTG 552
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        480 CATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTG 539

Qy        553 TGAGCAAGGACCCCCCGGTR 572
              ||||||||||||||||||:
Db        540 TGAGCAAGGACCCCCCGGTA 559


RESULT 105
DA858973
LOCUS       DA858973                 557 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA858973 PLACE7 Homo sapiens cDNA clone PLACE7010734 5', mRNA
            sequence.
ACCESSION   DA858973
VERSION     DA858973.1  GI:82362054
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7010734"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             7.1%;  Score 529.7;  DB 1;  Length 557;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  518;  Conservative   37;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       3701 CAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGT 3760
              ||||||||:|||||||||||:||||||||||||||||||:|||||||||:|||||
Db          1 CAGACAACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGT 60

Qy       3761 GCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAY 3820
              |||||||||||||||||||||||:| ||||||||||||||||||||||||||:||:
```

```
Db         61 GCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAAC 120

Qy       3821 ATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTC 3880
              |||||||||||||:|||||||||||||||||||||||||||||||||||:||||:||||
Db        121 ATAACTGTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTC 180

Qy       3881 CYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 3940
              |:||||||||||:|||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 239

Qy       3941 CTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
              ||||||||||:||||||||||||||||||||||||||:|||||||||||||||||||||
Db        240 CTTCGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTC 299

Qy       4001 ACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTA 4060
              ||||||||||||||||:||||:||||||||||||||:|||:|:|||||||||||||||
Db        300 ACTCCAGGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTA 359

Qy       4061 GGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
              |||||||:||||||||:|||||||||||||||||||:|:||||||:|||||||| |||||
Db        360 GGAGAAGAAAAAAGGGCAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCAT 419

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMA 4180
              |||||||::||||||||:|:|||||||||||||||||||||:|||:|||||||||||||:|
Db        420 GCCCATGCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAA 479

Qy       4181 CATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTC 4240
              ||||||||||||||||||:|||||||:||||:|||||||||||||:|||||||||:||||
Db        480 CATCAGGAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTC 539

Qy       4241 TTACACTGCYGGGGTCAT 4258
              |||||||||:||||||||
Db        540 TTACACTGCCGGGGTCAT 557
```

RESULT 106
DB381575/c
LOCUS DB381575 541 bp mRNA linear EST 03-DEC-2005
DEFINITION DB381575 PLACE4 Homo sapiens cDNA clone PLACE4000427 3', mRNA sequence.
ACCESSION DB381575
VERSION DB381575.1 GI:83094437
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 541)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute (HRI); cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and HRI.
FEATURES Location/Qualifiers
source 1. .541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000427"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"

Query Match 7.0%; Score 524.9; DB 1; Length 541;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 530; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

```
Qy       7040 TGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTA 7099
              |||||||||||||||||||||||  |||||||||||||||||||||||||||||||| ||
Db        541 TGGAGCCCAAGATGCAGTCCAAG-NTAAGATCTACCGCAGACCCCTGGACCGGCCTGNTA 483

Qy       7100 GCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCAC 7159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCAC 423

Qy       7160 AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCC 7219
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        422 AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCC 363

Qy       7220 CAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTT 7279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTT 303

Qy       7280 CCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAAC 7339
              ||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||||
Db        302 CCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAAC 243

Qy       7340 ACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTA 7399
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        242 ACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTA 183

Qy       7400 GGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGG 7459
              ||||||||||||||||||||||||||||||||||||:|||||:|||||||||||||||||
Db        182 GGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGG 123

Qy       7460 GACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTC 7519
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        122 GACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTC 63

Qy       7520 CCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCT 7579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db         62 CCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACT 3

Qy       7580 GC 7581
              ||
Db          2 GC 1
```

RESULT 107
AQ343877/c
LOCUS AQ343877 559 bp DNA linear GSS 07-MAY-1999
DEFINITION RPCI11-115K15.TV RPCI-11 Homo sapiens genomic clone RPCI-11-115K15, genomic survey sequence.
ACCESSION AQ343877
VERSION AQ343877.1 GI:4168773
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 559)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI11-115K15.TJ

```
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7544030"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-115K15"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 559;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  517;  Conservative   33;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       3591 AGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGAC 3650
             ||||||::|||||||||||||||||||||||||||||:||||||||||||||||||||
Db        559 AGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGAC 500

Qy       3651 ATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYG 3710
             |||||||||:|||||::||||||||||||||||||||||||||||||||||||||||:|
Db        499 ATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTG 440

Qy       3711 TTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGC 3770
             ||||||||||:|||||||||||||||||:|||||||||||:|||||||||||||||||
Db        439 TTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGC 380

Qy       3771 AACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCA 3830
             ||||||||||||||:| ||||||||||||||||||||||||||||:||:|||||||||||
Db        379 AACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACTGTCA 320

Qy       3831 ACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGA 3890
             ||||:||||||||||||||||||||||||||||||||||||:||||:||||||:||||||||
Db        319 ACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGACTGA 260

Qy       3891 TCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAG 3950
             |||:|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 TCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAG 201

Qy       3951 YGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAA 4010
             :||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        200 TGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAA 141

Qy       4011 CTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAA 4070
             ||||||||:||||:|||||||||||||||||:|||:|:|||||||||||||||||||:||
Db        140 CTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAA 81

Qy       4071 AAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRG 4130
             ||||||:|||||||||||||||||:|:|||||||:|||||||| ||||||||||||::|
Db         80 AAAGGGCAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAG 21

Qy       4131 CAATATGSARAGAAAGG 4147
             ||||||| |:  || ||
Db         20 CAATATGAAAGAAAGGG 4
```

```
RESULT 108
DA838102
LOCUS       DA838102                552 bp    mRNA    linear   EST 09-NOV-2005
DEFINITION  DA838102 PLACE6 Homo sapiens cDNA clone PLACE6004294 5', mRNA
            sequence.
ACCESSION   DA838102
VERSION     DA838102.1  GI:81277632
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .552
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6004294"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             6.9%;  Score 520.5;  DB 1;  Length 552;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  528;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 76

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 136

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 196

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        197 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 256

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        257 CGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 316

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
```

```
                 |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        317 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 376

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
                 |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        377 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 436

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
                 |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        437 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 495

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACC 537
                 ||:|||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        496 GAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACC 552


RESULT 109
DA375140
LOCUS       DA375140                563 bp    mRNA    linear   EST 07-NOV-2005
DEFINITION  DA375140 BRTHA2 Homo sapiens cDNA clone BRTHA2006318 5', mRNA
            sequence.
ACCESSION   DA375140
VERSION     DA375140.1  GI:81109189
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRTHA2006318"
                     /tissue_type="thalamus"
                     /clone_lib="BRTHA2"
                     /note="Vector: pME18SFL3"

  Query Match             6.9%;  Score 517;  DB 1;  Length 563;
  Best Local Similarity   94.5%;  Pred. No. 0;
  Matches  532;  Conservative    3;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       5787 CCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGG 5846
                 |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGCAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGG 60

Qy       5847 AAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTT 5906
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTT 120

Qy       5907 CACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGT 5966
                 |||||| ||  |||||||||||||||||||| |||||| |||||||||||||||||| ||
Db        121 CACCCATACCAGTATGTCTGATGGGGGTGGAATTCAAGGTCAGGCAAGAGAAAAACAAGT 180

Qy       5967 AAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACT 6026
                 ||| |||| |||||||||||| |||:||| |||| |||||| ||||||||||||||||||
Db        181 AAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCCTAGCCCCTACAAAGGACT 240

Qy       6027 AGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAA 6086
                 || ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        241 AGTTCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTGAGCCTATTTAA 300

Qy       6087 TACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATG 6146
                 |||||||||||| |||||||||||||||| |||||||||||||||||||||||||| ||
Db        301 TACCACCCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCTACTAACTGTTGGATGTG 360

Qy       6147 CCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTT 6206
                 |||||||||| ||||||:|||||| |||||||||||| ||||||||||||||||||||
Db        361 CCTCCCCCTGCACTTCAGGCCATACATTTCAATCCCTGTTCCTGAACAATGGAACAACTT 420

Qy       6207 CAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAAT 6266
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        421 CAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAAT 480

Qy       6267 AACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTC 6326
                 |||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||| |||
Db        481 AACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAACCAGCTC 540

Qy       6327 CCAATGCATCAGGTGGGTAACTC 6349
                 |||||||||||||||||||| |
Db        541 CCAATGCATCAGGTGGGTAACAC 563


RESULT 110
CR735592
LOCUS       CR735592                554 bp    mRNA    linear   EST 27-AUG-2004
DEFINITION  CR735592 NIH_MGC_21 Homo sapiens cDNA clone IMAGp958M21530 ;
            IMAGE:3840572 5', mRNA sequence.
ACCESSION   CR735592
VERSION     CR735592.1  GI:51584854
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E.,
            Peters,M., Radelof,U. and Schneider,D.
  TITLE     I.M.A.G.E. cDNA Clone Collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; IMAGp958M21530.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

/db_xref="taxon:9606"
/clone="IMAGp958M21530 ; IMAGE:3840572"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

Query Match 6.9%; Score 515; DB 1; Length 554;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 512; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6282 CCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTG 6341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 CCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTG 97

Qy       6342 GGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTAC 6401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 GGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTAC 157

Qy       6402 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGT 6461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGT 217

Qy       6462 GCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCG 6521
             |||||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCG 277

Qy       6522 CAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGG 6581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGG 337

Qy       6582 TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACT 6641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACT 397

Qy       6642 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 6701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 457

Qy       6702 CCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGG 6761
             ||||||||||||||||||:||||||||||||||||:||||||||||||:|||||
Db        458 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGG 517

Qy       6762 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTAT 6798
             |||||||||||||||||||||||||||||||||||||
Db        518 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTAT 554
```

RESULT 111
AQ193544/c
LOCUS AQ193544 594 bp DNA linear GSS 16-SEP-1998
DEFINITION CIT-HSP-2381P7.TR CIT-HSP Homo sapiens genomic clone 2381P7, genomic survey sequence.
ACCESSION AQ193544
VERSION AQ193544.1 GI:3603659
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2381P7.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .594
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2381P7"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"

Query Match 6.8%; Score 507.7; DB 1; Length 594;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 508; Conservative 45; Mismatches 28; Indels 1; Gaps 1;

```
Qy       2814 ATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCA 2873
             |||||||||||||||||||||||| |||||||||||||||||| |||| |||| |:
Db        594 ATCCCTGTACATCCTGACTCTCAAATCTTGTTTGCCTTTGAAGATCCTTCGAACCTAATG 535

Qy       2874 TCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGC 2933
             ||||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        534 TCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGC 475

Qy       2934 CAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAK 2993
             |||||| |||||||||||||||:||:|::|||||||||||||||||||||||||||:|||
Db        474 CAGGCACTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCGGTAC 415

Qy       2994 GTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCR 3053
             |||||||||||| ||||:||:|||::|||| ||||||||||||||||||||||||||||:
Db        414 ATGGATGATTTATTTTTAGCTGCCCGTTCAAAAACCTTGTGCCATCAAGCCACCCAAGCG 355

Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:|||||||||||:||||||||| |||:|||||||||||:|:|:||||||:||||||
Db        354 CTCTTAAATTTCCTCGCTACCTGTGGCCACAAGGTTTCCAAACCAAAGGCTCAGCTCTGC 295

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||:||||:||||||||||||||||:|||||||||||||||
Db        294 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAG 235

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
             |||::: |||||||||||||||||||||||||||:|||||||||||||||||||||:|::
Db        234 GAACGTTTCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGG 175

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
             |||||||||:|||:|||:|||||||||||:||||||| |||||| :  | :||||||||||
Db        174 TTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT-CCCAGGCACAGTGAAATAGCCAG 116

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
             |:||||||:|||||:|||||| |||||||||||| | || ||||||||||:|||||||||:
Db        115 GCCATTATATACACTAATTATGGAAACTCAGACAACCTATACCCATTTAGTAAGATGGAC 56

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCT 3395
             |:||||||:::||| |||||||||||||  |   ||| |||
Db         55 ACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAACCCAAGCCCT 14
```

RESULT 112

```
DB028309
LOCUS       DB028309                534 bp    mRNA    linear   EST 08-NOV-2005
DEFINITION  DB028309 TESTI2 Homo sapiens cDNA clone TESTI2011169 5', mRNA
            sequence.
ACCESSION   DB028309
VERSION     DB028309.1  GI:81209045
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI2011169"
                     /tissue_type="testis"
                     /clone_lib="TESTI2"
                     /note="Vector: pME18SFL3"

  Query Match             6.7%;  Score 501.7;  DB 1;  Length 534;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  502;  Conservative   23;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       4084 ATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGA 4143
              ||||||||||||||:|:||||||:||||||| ||||||||||||::||||||||:|:|||
Db          1 ATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAAAGA 60

Qy       4144 AAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATT 4203
              ||||||||||||||||:|||:||||||||||||:||||||||||||||||:|||
Db         61 AAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTAGGAAATT 120

Qy       4204 ATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAA 4263
              ||||:||||:|||||||||||||:||||||||:|||||||||||||:||||||||| ||
Db        121 ATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAA 180

Qy       4264 AGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGC 4323
              |||||||:||||||||||:|||:|||:||||||||:||||||||:|||||||||||||||
Db        181 AGGAAAGGAAAGGGAAATAGAAGAGAACTGCCAAGCAGATATTGAAGCCAAAAGAGCTGC 240

Qy       4324 AAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTT 4383
              |||||||||||||||||||||||||||||| ||||  ||| ||||||||||||||||| |
Db        241 AAGGCAGGACCCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCCT 300

Qy       4384 CCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG-CA 4442
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
```

```
Db        301 CCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGGACA 360

Qy       4443 GTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTA 4502
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 GTTTTCTCCCCTCGGGACGGCTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTA 420

Qy       4503 TCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCC 4562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCC 480

Qy       4563 ATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 534


RESULT 113
BG000981
LOCUS       BG000981                503 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  IL0-GN0246-131100-501-c06 GN0246 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG000981
VERSION     BG000981.1  GI:12438850
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL0&t2=IL0-GN0246-
            131100-501-c06&t3=2000-11-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0246"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             6.7%;  Score 501;  DB 1;  Length 503;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches  498;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
              |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 60

Qy       5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 120

Qy       5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 180

Qy       5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db        181 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 240

Qy       5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 300

Qy       5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 360

Qy       5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 420

Qy       5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 480

Qy       5555 AAAGTAACTAAAATCATAAATCC 5577
              |||||||||||||||||||||||
Db        481 AAAGTAACTAAAATCATAAATCC 503


RESULT 114
BG012022/c
LOCUS       BG012022                 531 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  QV3-GN0335-131200-541-d12 GN0335 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG012022
VERSION     BG012022.1  GI:12460813
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-GN0335-
            131200-541-d12&t3=2000-12-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 530.

FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0335"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             6.7%;  Score 500.3;  DB 1;  Length 531;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  514;  Conservative   15;  Mismatches    2;  Indels    3;  Gaps    2;

Qy       4531 TCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGG 4590
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        531 TCAAACCTTTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGG 472

Qy       4591 ACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAA 4650
              ||||||||||||||||||||||||||||:|||:|||||||||||||:|||||||:|:|||
Db        471 ACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAA 412

Qy       4651 TAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGR 4710
              ||||||||||||:||||||||||||||||||||:||||||:||||||||||||||||||:
Db        411 TAATCCCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGG 352

Qy       4711 ARAARACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTA 4770
              |:||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGAAGACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTA 292

Qy       4771 GTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCC 4830
              |||||||||:||||||||||||||||||:||||||||||||||||||||||||||||||
Db        291 GTCTGGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCC 232

Qy       4831 CAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGC 4890
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        231 CAAGAGGTAATAAAGGCACTAGTTCATGAAATAATT-CCAGATTCGGACTTCCCCGAGGC 173

Qy       4891 TTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCG 4950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCG 113

Qy       4951 TTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAG 5010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 TTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAG 53

Qy       5011 AAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCAC 5064
              |||||||||||||:|||||||||||||||  |||||||||||||||| |||| |
Db          52 AAAATGAATGAAACACTCAAAGGACATCT--AAAAGCAAACCCAGGACACCCGC 1


RESULT 115
DA848644
LOCUS       DA848644                 578 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA848644 PLACE6 Homo sapiens cDNA clone PLACE6017757 5', mRNA
            sequence.
ACCESSION   DA848644
VERSION     DA848644.1  GI:82111120
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
```

```
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6017757"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             6.6%;  Score 496.9;  DB 1;  Length 578;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches  502;  Conservative   10;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db         66 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 125

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 185

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 245

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        246 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 304

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
              ||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        305 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 364

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
              ||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        365 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 424

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        425 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 484

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 544

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGT 5865
              |||||||||||||||||||||||||||| |||||
Db        545 TCATTATTGGACAGGAAAAATGATTAATTCTAGT 578


RESULT 116
AQ012843
LOCUS       AQ012843                 558 bp    DNA     linear   GSS 06-JUN-1998
DEFINITION  CIT-HSP-2300M15.TF CIT-HSP Homo sapiens genomic clone 2300M15,
            genomic survey sequence.
ACCESSION   AQ012843
VERSION     AQ012843.1  GI:3185408
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2300M15.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7154635"
                     /db_xref="taxon:9606"
                     /clone="2300M15"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             6.6%;  Score 491.7;  DB 1;  Length 558;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  498;  Conservative   33;  Mismatches   27;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||||||:|:|||||||||: ||||| ||  |||||||||||::|||| ||||||
Db          1 GCAAAGGGTTGGCCTCATTGTTTACAGGTAGGGGCAGCAGTAGCAGTCTTAGTTTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |||||||||||||||||:|||||||||||||||||||||||||||:|||||::||
Db         61 ACAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              |||||||||||||||| ||||||||||||||||||||: ||||||||||: |||| |
Db         121 ATACTCACTGCTAAAGAAGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              ||||||||||: ||||||||||:|||| || ||||||||||||||||||:| ||
Db         181 CTATTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              ||||||||||||||||||||||||||:||:|||||||||||||:||||| ||||||||
Db         241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
              |||||  |||||||||||||:|||||:|||||:||||||||||:||||| ||||||||
Db         301 TATGCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACCT-CAACTTGTA 359
```

```
Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             |||||||||||||||||| | |||||||||||| ||||||:|||||||||||| ||||
Db        360 TACTGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAG 419

Qy       3970 TGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGA 4029
             |||||||||||||:|||||||||||| |||||||||||||||||||||:| ||:|||||
Db        420 TGATAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTTACCTGGCAGA 479

Qy       4030 ACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATAC 4089
             ||||||||||:| |:|:|||||||||||||||||||:||||||||: |||||||| |
Db        480 ACTAATAGCCCTTACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTTAATATATATTC 539

Qy       4090 AGACTCTRARTATGCTYAC 4108
             |||||||:|:||||||:||
Db        540 AGACTCTAAGTATGCTTAC 558


RESULT 117
BX492789
LOCUS       BX492789                 528 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp781H0323_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781H0323 5', mRNA sequence.
ACCESSION   BX492789
VERSION     BX492789.1  GI:32004620
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781H0323) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781H0323"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"

  Query Match             6.5%;  Score 488.5;  DB 1;  Length 528;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  487;  Conservative   29;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       3977 GGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATA 4036
             ||||||:||||||||||||||||||||||||||||||||||:||||:||||||||||||
Db          1 GGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTAGCAGAACTAATA 60

Qy       4037 GCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCT 4096
             |||:|||:|:|||| ||||||||||||||||:||||||||:||||||||||||||||||
Db         61 GCCCTCACTTGGGCGCTAGAATTAGGAGAAGAAAAAAGGGCAAATATATATACAGACTCT 120

Qy       4097 RARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTA 4156
                :|:||||||:|||||||| ||||||||||||::||||||||:|:||||||||||||||||
Db         121 AAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAAAGAAAGGGAATTCCTA 180

Qy       4157 ACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTA 4216
             |||||:|||:||||||||||||:||||||||||||||||||||:|||||||:||||:|||
Db         181 ACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGGCTGTA 240

Qy       4217 CAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGG 4276
             ||||||||||:|||||||:||||||||||||:|||||||| |||||||||:|||||
Db         241 CAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAAGGAAAGGAAAGG 300

Qy       4277 GAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCT 4336
             |||||| |||:|||:|||||||||:||||||||||:|||||||||||||||||||||||||
Db         301 GAAATATAAGAGAACTGCCAAGCAGATATTGAAGCCAAAAGAGCTGCAAGGCAGGACCCT 360

Qy       4337 CCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAG 4396
             |||||||||||||||| ||||  ||| ||||||||||||||||| ||||||||||||||
Db         361 CCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCCTCCGGGAAACCAAG 420

Qy       4397 CCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG-CAGTTTTCTCCCCTC 4455
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         421 CCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGGACAGTTTTCTCCCCTT 480

Qy       4456 GGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||||||| ||||||||||||||||||||||||||||||||||||||||
Db         481 GGGACGGCTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 528


RESULT 118
BF996761
LOCUS       BF996761                 610 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  MR1-GN0173-071100-009-b06 GN0173 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF996761
VERSION     BF996761.1  GI:12403084
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-GN0173-
            071100-009-b06&t3=2000-11-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 610.
FEATURES             Location/Qualifiers
     source          1. .610
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0173"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
```

SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 6.5%; Score 488; DB 1; Length 610;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 505; Conservative 57; Mismatches 32; Indels 4; Gaps 3;

```
Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTT 3133
Db         17 CCGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACGT 76

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACT 3193
Db         77 AGGGCTAAAATTATCCAAAGGCACCAGGTCCTTCAGTAAGGAACGTATCCAGCCTATACT 136

Qy       3194 GGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
Db        137 GGCTTATCCCCATCCCAAAACCCTAAAGCAACTACGAGGGTTCCTTGGCATAACACGCTT 196

Qy       3254 CTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTA 3313
Db        197 CTGCCAAATATGGATT-CCCAGGTACGGCAAAATAGCCAGGCTATTATATACACTAATTA 255

Qy       3314 AGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTT 3373
Db        256 AGGAAACTCAGAAAGCCAATACCCATTTAGTAAGAAGGACACCTGAAGCAGAAGCGGCTT 315

Qy       3374 TCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCA 3433
Db        316 TCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTATTAAGCTTGCCAACGGAGCG 373

Qy       3434 AGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGR 3493
Db        374 AGACTTTTCTTTATATGTCACAG-AAAAAAGAGGAATAGCTCTAGGAATCCTTACACAGG 432

Qy       3494 TCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAA 3553
Db        433 TCCGAGAGACCAGCTTGCAACCCATGGCTTACCTAAGTAAGGAAATTGATGTAGTGGCAA 492

Qy       3554 AGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAG 3613
Db        493 AAGGTTGGCCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAG 552

Qy       3614 TTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
Db        553 TTAAAATAATACAGGGAAGAGATCTTACCGTGTGGACATCTCATGATGTTGACAGCAT 610
```

RESULT 119
DB066140
LOCUS DB066140 545 bp mRNA linear EST 11-DEC-2005
DEFINITION DB066140 TESTI4 Homo sapiens cDNA clone TESTI4007381 5', mRNA sequence.
ACCESSION DB066140
VERSION DB066140.1 GI:83545280
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
FEATURES Location/Qualifiers
source 1. .545
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI4007381"
/tissue_type="testis"
/clone_lib="TESTI4"
/note="Vector: pME18SFL3"

Query Match 6.5%; Score 484; DB 1; Length 545;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 492; Conservative 51; Mismatches 1; Indels 16; Gaps 3;

```
Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
Db          1 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 59

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
Db         60 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 119

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
Db        120 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGC 165

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
Db        166 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACAG 224

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
Db        225 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCATACC 284

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
Db        285 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 344

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
Db        345 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 404

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
Db        405 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 464

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
Db        465 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 524

Qy       3767 GTGCAACTCTTAACCCAGYC 3786
Db        525 GTGCAACTCTTAACCCAGCC 544
```

RESULT 120

```
DB381433/c
LOCUS       DB381433                537 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB381433 PLACE4 Homo sapiens cDNA clone PLACE4000151 3', mRNA
            sequence.
ACCESSION   DB381433
VERSION     DB381433.1  GI:83093695
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000151"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             6.4%;  Score 482.3;  DB 1;  Length 537;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  491;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       7070 TCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAG 7129
              | | ||||| |||||||||||||||||| |||||| |||||||||||||||| ||||||||
Db        511 TTTTCCGCAAACCCCTGGACCGGCCTG-TAGCCCNCGATCTGATGTTAATGCCATCAAAG 453

Qy       7130 GCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAA 7189
              |||||| |||||||||||||||  |||||||| | || ||||||||||||||||||||||
Db        452 GCACCCNTCCTGAGGAAATCTCAAGTGCACAACNTNTANTACGCCCCAATTCAGCAGGAA 393

Qy       7190 GCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATG 7249
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATG 333

Qy       7250 GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAG 7309
              |||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||
Db        332 GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAG 273

Qy       7310 STGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGAC 7369
              :||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        272 CTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGAC 213

Qy       7370 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAA 7429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAA 153

Qy       7430 TCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTT 7489
              ||||:||||||:||||||||||||||||||||||||||||||||||||||||||||:||
Db        152 TCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTT 93

Qy       7490 CGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 CGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 33

Qy       7550 TGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||||||||||||||||:||||
Db          32 TGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 121
AA781423/c
LOCUS       AA781423                494 bp    mRNA    linear   EST 31-DEC-1998
DEFINITION  aj26c03.s1 Soares_testis_NHT Homo sapiens cDNA clone 1391428 3'
            similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AA781423
VERSION     AA781423.1  GI:2840754
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1645   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="1391428"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares_testis_NHT"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from mRNA obtained from
                     Clontech Laboratories, Inc., and primed with  a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "

  Query Match             6.4%;  Score 480.4;  DB 1;  Length 494;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  475;  Conservative   11;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7094 CTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAG 7153
              |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 CTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAG 435

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
```

```
Db        434 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAA 375

Qy       7214 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 7273
Db        374 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 315

Qy       7274 GGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCT 7333
Db        314 GGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCT 255

Qy       7334 TTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 7393
Db        254 TTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 195

Qy       7394 TAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGC 7453
Db        194 TAACTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGC 135

Qy       7454 AGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTT 7513
Db        134 AGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACNCCCTT 75

Qy       7514 TGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTT 7573
Db          74 TGGGTCCCCTCCCTTTGTGTGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTT 15

Qy       7574 GCARCTGCR 7582
Db          14 GCAACTGCA 6
```

```
RESULT 122
AA613238/c
LOCUS       AA613238                578 bp    mRNA    linear   EST 03-OCT-1997
DEFINITION  no19f04.s1 NCI_CGAP_Phe1 Homo sapiens cDNA clone IMAGE:1101151 3'
            similar to SW:POL_SMSAV P03359 POL POLYPROTEIN ;, mRNA sequence.
ACCESSION   AA613238
VERSION     AA613238.1  GI:2464276
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1101151"
                     /tissue_type="pheochromocytoma"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_Phe1"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Pheochromocytoma. 5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                     Average insert size:  1.3 kb."
```

```
  Query Match             6.4%;  Score 480.4;  DB 1;  Length 578;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  526;  Conservative    5;  Mismatches   46;  Indels    2;  Gaps    2;

Qy       4860 AATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCA 4919
Db        578 AATTATTCCCAGATGC-GACTTCCCCGAGGCTTACAGAGTGACAATGGCCCTGCTTTCAA 520

Qy       4920 GGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGC 4979
Db        519 GTCAACAGTAACCCAGGGACTATCGCAGATGTCAGGTATACAATATCACTTACACTGCAC 460

Qy       4980 CTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCT 5039
Db        459 CTGGAGTCCACAATCCTCGGGGAAGGTCGAGAAAATGAACGAAACACTCAAACAACATCT 400

Qy       5040 AAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAA 5099
Db        399 AAAAAAGCTAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTACT 340

Qy       5100 AAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5159
Db        339 AAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGACG 280

Qy       5160 GCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACAT 5218
Db        279 GCCCTTCCTAACCAATGACCTTGTGCTTGACCGAGAGACGGCCAACTTAGTTGCAGATAT 220

Qy       5219 CACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGA 5278
Db        219 CACCTCCTTAGCCAAATATCAGCAAGTTCTTAAAACATGACAAGAAGCCTGTCCCTGAGA 160

Qy       5279 AGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAA 5338
Db        159 AGAGGGAAAGGAATTATTCCACCCTGGTGACATGGTATTAGTCAAGCCCCTTCCTTCTGA 100

Qy       5339 TTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAAC 5398
Db         99 TTCCCCATCCCTAGATACATCCTGGGAAGGACGCTACCCGGTCATTTTATCTACCCCAAC 40

Qy       5399 TGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
Db         39 CGCGGTTAAAGTGGCTGGAGTGGAGTCTAGGATACATCA 1
```

```
RESULT 123
B59017/c
LOCUS       B59017                  595 bp    DNA     linear   GSS 20-JUN-1998
DEFINITION  CIT-HSP-2014K19.TR CIT-HSP Homo sapiens genomic clone 2014K19,
            genomic survey sequence.
ACCESSION   B59017
VERSION     B59017.1  GI:2613735
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2014K19.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
```

```
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7042821"
                     /db_xref="taxon:9606"
                     /clone="2014K19"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             6.4%;  Score 480.2;  DB 1;  Length 595;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  511;  Conservative   43;  Mismatches   41;  Indels    4;  Gaps    3;

Qy       2824 ATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCA 2883
              ||||||||||||||||||||||||||||||||||  ||||||||||:|||||||||||
Db        595 ATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCATTCAAACCCAACATCTCAACTCA 536

Qy       2884 CCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAG 2943
              |||||| |:|||||||||||||||||||||||||:|||||| |||||||||||||| |||
Db        535 CCTGGATTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATGTATTTGGCCAGGCACTAG 476

Qy       2944 CCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATT 3003
              | |||||||||||:||:|::|||||||||||||||| | ||||||:| |  |||||||||
Db        475 CTCAAGACTTGAGCCAGTTCTCATACCTGGACACTCCTCTCCTTCAGCACATGGATGATT 416

Qy       3004 TACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATT 3063
              |||||||:||:|||: ||||||||||||||||||||||||||||||: |||| ||||
Db        415 TACTTTTAGCTGCCCTTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCATTCTTTAATT 356

Qy       3064 TCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGG 3123
              ||||| |:|||||| | || |:|||||||||||| |:|:| |||:|||||||||||| |||
Db        355 TCCTCTCCACCTGT-GATATAAGGTTTCCAAACAAAAGGATCAGCTCTGCTCACAGTAGG 297

Qy       3124 TTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCC 3183
              |||||||||||| :||||:|||||||||||||||:|||||||||||||||||:::||||
Db        296 TTAAATACTTAGTGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCC 237

Qy       3184 AGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCR 3243
              ||||||| ||||||| |||||||||:|||||| |||||||||||||:| :|||||||||:
Db        236 AGCCTATCCTGGCTTTTCCTCATCCCAAAACCGTAAAGCAACTAAGAGCGTTCCTTGGCA 177

Qy       3244 TAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWAT 3303
              |||:|||:|||||| |||:|| |||| ||| |||:  ||:|||||||||| :|||||:||
Db        176 TAACAGGCTTCTGCTGAATATAGATT-CCCTGGTACAGCAAAATAGCCAGATCATTATAT 118

Qy       3304 ACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYM 3363
              |||:|||||||||||||||||| ||||||||||||:|||||||:|:||||||::
Db        117 ACACTAATTAAGGAAACTCAGAAAGGCAATACCCATTTAGTAAGATGGACACCTGAAGCA 58

Qy       3364 RAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTG 3422
              :||||| ||||||||||||  ||  | | |||  |||| |||: ||||||||||:|||
Db          57 GAAGTGACTTTCCAGGCCCTAAAGAAGGCCCT--AACCTAAGCTCCAGTGTTAAGCTTG 1


RESULT 124
BG573364
LOCUS       BG573364                631 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602595137F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4724433 5',
            mRNA sequence.
ACCESSION   BG573364
VERSION     BG573364.1  GI:13581017
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1585  row: i  column: 10
            High quality sequence stop: 628.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4724433"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_79"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); 5' and 3' adaptors were used in cloning
                     as follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3'
                     and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.3
                     kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."

  Query Match             6.4%;  Score 480;  DB 1;  Length 631;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  537;  Conservative   11;  Mismatches    4;  Indels    6;  Gaps    6;

Qy         47 CCCCCTTTGGGTCCCTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAA 106
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CCCCCCTTGGGTCCCTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAA 66

Qy        107 ATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCA 166
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         67 ATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGC-TTTGCTCA 125

Qy        167 CCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGA 226
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        126 CCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGA 185

Qy        227 TCCTGCAGGGTGTCCGCTGTGCTCC-TGATCCAGCGARGCGCCCATTGCCGCTCCCAATT 285
              |||||||||||||||||||||||| ||||||||||:||||||||||||||||||||||||
Db        186 TCCTGCAGGGTGTCCGCTGTGCTCCGTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATT 245

Qy        286 GGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGC 345
              |||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db        246 GGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGG-TTGTTCTAATTGAGC 304

Qy        346 TGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACT 405
              |||||||| ||||||||||||||||||||||||||||||||||||||||||:||||
Db        305 TGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACT 364

Qy        406 ATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTC 465
              |||||||||||||||||||||||||||||||||||||||||||||||||:||| ||||||
Db        365 ATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTC 423
```

```
Qy        466 CAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAG 525
              ||||||||||||||||:||||||||||||||||||||| ||||||||:|||||||||
Db        424 CAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGG-CTGCTACCATCTTGGAAG 482

Qy        526 TGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRAC 585
              ||||||||||||||||||||||||||||||||||||||||||:||| |||||||:||
Db        483 TGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACA-TTTGGCAAC 541

Qy        586 CAMSRACGGACATCCMAA 603
              ||:::|||||||||||:||
Db        542 CACGAACGGACATCCAAA 559


RESULT 125
DB381156/c
LOCUS       DB381156                544 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB381156 PLACE3 Homo sapiens cDNA clone PLACE3000098 3', mRNA
            sequence.
ACCESSION   DB381156
VERSION     DB381156.1  GI:83091618
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE3000098"
                     /tissue_type="placenta"
                     /clone_lib="PLACE3"
                     /note="Vector: pME18SFL3"

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 544;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  504;  Conservative    9;  Mismatches   21;  Indels    3;  Gaps    2;

Qy       7046 CCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCAC 7105
              ||||||| |||||||||  | |  | |  |     |||| ||||||||| || |||||| |
Db        535 CCAAGATNCAGTCCAAG--TTANGTTTTNCCGCAACCCTTGGACCGGCTTGNTAGCCCCC 478

Qy       7106 GATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTC 7165
              |||||||||||||| |||||||||||||| || |||||||||||||||||||||||| |
Db        477 GATCTGATGTTAATGCCATCAAAGGCACCCNTCNTGAGGAAATCTCAGCTGCACAACTTT 418

Qy       7166 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTC-GGCCAACCTCCCCAACA 7224
              || ||||||||||||||||||||||||||||||||||:||| |||||||||||||||||
Db        417 TATTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGGCCAACCTCCCCAACA 358

Qy       7225 GCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAG 7284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAG 298

Qy       7285 GCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGG 7344
              |||||:|||||||||:|||||||||:||||||||||||||||||||||||||||||||||
Db        297 GCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGG 238

Qy       7345 GCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAA 7404
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAA 178

Qy       7405 AGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAA 7464
              ||||||||||||||||||||||||||||||:||||||:|||||||||||||||||||||||
Db        177 AGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAA 118

Qy       7465 TGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTC 7524
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        117 TGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTC 58

Qy       7525 CCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          57 CCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 126
DB068316
LOCUS       DB068316                544 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB068316 TESTI4 Homo sapiens cDNA clone TESTI4010208 5', mRNA
            sequence.
ACCESSION   DB068316
VERSION     DB068316.1  GI:83100278
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI4010208"
                     /tissue_type="testis"
```

/clone_lib="TESTI4"
/note="Vector: pME18SFL3"

Query Match 6.3%; Score 474.5; DB 1; Length 544;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 493; Conservative 31; Mismatches 6; Indels 7; Gaps 3;

```
Qy       2670 GGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTA 2729
             ||||||||||||||||||||||||||||| ||||| ||||| | | |||| |||||||||
Db         15 GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGCTGTTGTTCCTCTATAGCCAGCTGTA 73

Qy       2730 CCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTG 2789
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||:|||||
Db         74 CCTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTG 133

Qy       2790 GACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 2849
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GACCTTCAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 193

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
             ||||||||||||||||||||:||||||||||||||||:|||||||||||||||||||||
Db        194 TTTGAAGATACTTCAAACCCAACATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTC 253

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
             ||||||||:|||||||||||||||||||||||||||||||||||||:||:|::||||||
Db        254 AGGGATAGTCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATAC 313

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
             |||||||  ||||||||||:|||:|||||||||||||||:||:|||::|||||||||||
Db        314 CTGGACA--CTTGTCCTTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACC 371

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
             ||||||||||||||||||||||:|||||:||||||||||:|||||||||||||:||||
Db        372 TTGTGCCATCAAGCCACCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTT 431

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||||||||||    ||||||||:||||:|||||||
Db        432 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 487

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||
Db        488 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCAT 544
```

RESULT 127
DB096981
LOCUS DB096981 569 bp mRNA linear EST 10-DEC-2005
DEFINITION DB096981 TESTI4 Homo sapiens cDNA clone TESTI4048110 5', mRNA sequence.
ACCESSION DB096981
VERSION DB096981.1 GI:83478258
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
FEATURES Location/Qualifiers
source 1. .569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI4048110"
/tissue_type="testis"
/clone_lib="TESTI4"
/note="Vector: pME18SFL3"

Query Match 6.3%; Score 474.3; DB 1; Length 569;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 524; Conservative 19; Mismatches 26; Indels 11; Gaps 4;

```
Qy       2481 CAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCC 2540
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 CAAGTTAAACTAAAGGATTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCC 60

Qy       2541 AAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAA 2600
             |||||||||||||||| |||||||||||||||||| |||||||||||||| ||||||||
Db         61 AAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAA 120

Qy       2601 CCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGG 2660
             |||||||  |  |||||||||| | |         |||   |  | ||||||||||   |
Db         121 CCATGCAGTAACCCCTGCAGTACTCC-------AATTTTAGGAGTACAGAAACCCAACAG 173

Qy       2661 ACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATAC 2720
             |||||||| |||||||||||||||||||||||||||| ||||| ||||| | | ||||
Db         174 ACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGCTGTTGTTCCTCTATAG 231

Qy       2721 CCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTT 2780
             |||||||||||||||||||||||| ||::|||||||||||||||||||||||||||||||
Db         232 CCAGCTGTACCTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTT 291

Qy       2781 ACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTC 2840
             |||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         292 ACAGTCCTGGACCTTCAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTC 351

Qy       2841 TTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCC 2900
             |||||||||||||||||||||||||||||||:|||||||||||||||||||:|||||||
Db         352 TTGTTTGCCTTTGAAGATACTTCAAACCCAACATCTCAACTCACCTGGACTATTTTACCC 411

Qy       2901 CAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCAR 2960
             |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||:
Db         412 CAAGGGTTCAGGGATAGTCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAA 471

Qy       2961 TYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRT 3020
             |::|||||||||||||  ||||||||||:|||:|||||||||||||||||:||:|||::|
Db         472 TCCTCATACCTGGACA--CTTGTCCTTCGGTAGGTGGATGATTTACTTTTGGCCGCCCAT 529

Qy       3021 TCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMA 3060
             |||||||||||||||||||||||||||||||:|||||:|
Db         530 TCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTCA 569
```

RESULT 128
DB381477/c
LOCUS DB381477 526 bp mRNA linear EST 03-DEC-2005
DEFINITION DB381477 PLACE4 Homo sapiens cDNA clone PLACE4000228 3', mRNA sequence.
ACCESSION DB381477
VERSION DB381477.1 GI:83093739
KEYWORDS EST.
SOURCE Homo sapiens (human)

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 526)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000228"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             6.3%;  Score 472.1;  DB 1;  Length 526;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  493;  Conservative   10;  Mismatches   17;  Indels    3;  Gaps    2;

Qy       7059 CAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAA 7118
             | |   | || | | ||||| |||||||| ||||||||:    || |||| |||||||||
Db        520 CCAAGNTTAGNTNTCCCGCAAACCCCTGGNCCGGCCTGT--AGCCNCGATNTGATGTTAA 463

Qy       7119 TGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAA 7178
             |||||||||||| |||||||||||||||||||||||||||| || ||||||||||
Db        462 TGACATCAAAGGC-CCCCTCCTGAGGAAATCTCAGCTGCACAACCTNTANTACGCCCCAA 404

Qy       7179 TTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 7238
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        403 TTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 344

Qy       7239 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATC 7298
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        343 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATC 284

Qy       7299 CYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGY 7358
             |:||||||||:|||||||||||||||||||||||||||||||||||||||||||||||:
Db        283 CCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGC 224

Qy       7359 TCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAA 7418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 TCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAA 164

Qy       7419 GAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAA 7478
             ||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
Db        163 GAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAA 104

Qy       7479 ACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAG 7538
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 ACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTATATGGGAG 44
```

```
Qy       7539 CTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||||||||||||||||||||||||||||||:||||
Db         43 CTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 129
BQ365207/c
LOCUS       BQ365207                494 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC6-GN0070-220800-021-D12 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365207
VERSION     BQ365207.1  GI:21040719
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            220800-021-D12&t3=2000-08-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 40
            High quality sequence stop: 432.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             6.3%;  Score 471.4;  DB 1;  Length 494;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  469;  Conservative   10;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       6608 CTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCG 6667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 CTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCG 435

Qy       6668 ACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATC 6727
             |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        434 ACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATC 375

Qy       6728 GAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAG 6787
             |||||||||||||||:|||||||||||||:|||||||||||||||||||||||||||||
Db        374 GAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAG 315
```

```
Qy       6788 AATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAG 6847
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||
Db        314 AATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAG 255

Qy       6848 ATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCC 6907
             ||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db        254 ATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCC 195

Qy       6908 RATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCC 6967
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCC 135

Qy       6968 TCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTG 7027
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        134 TCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTG 75

Qy       7028 TRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGG 7087
              |:|||||||||||||| ||||||||||||||||||||||||   ||||||||||
Db         74 TAAAACTACAAATGGAGACCAAGATGCAGTCCAAGACTAAGATCTACGCGAGACCCCTGG 15

Qy       7088 ACCGG 7092
                |||
Db         14 GACGG 10
```

RESULT 130
BZ611361/c
LOCUS BZ611361 598 bp DNA linear GSS 08-JUN-2003
DEFINITION WHAAP03TF Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-4B6, genomic survey sequence.
ACCESSION BZ611361
VERSION BZ611361.1 GI:31519922
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .598
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-4B6"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

Query Match 6.3%; Score 470.1; DB 1; Length 598;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 504; Conservative 48; Mismatches 32; Indels 7; Gaps 4;

```
Qy       2833 CTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTR 2892
             | |||| | |||||||||| |||||| |||||||| ||:| |||||||||||||||||:
Db         587 CCCAATCCCTGTTTGCCTTGGAAGATCCTTCAAACTCAACGTCTCAACTCACCTGGACTG 528

Qy       2893 TTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACT 2952
              ||||| ||||||||||||||||:|||||||||||||||||||||||||||||||||| |
Db         527 -TTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGATT 469

Qy       2953 TGAGYCA---RTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTT 3009
              ||||:||   :|::|||||||||||||||||||||||||| :|||:|||||||||||||||
Db         468 TGAGCCAATTATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGATTTACTTT 409

Qy       3010 TRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCG 3069
              |:||: ||::|||||||||||||||||||||||||||||||| :|||||:|||||||||
Db         408 TAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCA 349

Qy       3070 CYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAAT 3129
              |:|||||||||||||:|||||||||||:|:|:|:|||||:|||||||||||||||||||||
Db         348 CCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAAT 289

Qy       3130 ACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTA 3189
              |||||||:||||:|||||||||||||    |:|||||||||||||||:::|||||||||||
Db         288 ACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACATCCAGCCTA 229

Qy       3190 TACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAG 3249
              ||||||||||||||||||:|||||||||||||||||||:|::|||||||||:|||:||
Db         228 TACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAG 169

Qy       3250 GYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTA 3309
              |:|||| |||||: |||||| |||||||:  ||:|||||| ||| :|||||:|||||:||
Db         168 GTTTCTACCGAATGTGGATT-CCCAGGTACAGCAAAATAGGCAGATCATTATATACACTA 110

Qy       3310 ATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTG 3369
              ||||| |||||||||||||||||||||||||||||:|||||||||:|:||||||:::|||
Db         109 ATTAAAGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCA 50

Qy       3370 GCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420
              |||||||||||||   ||  ||| |   ||||||||:|||||||||||||:|
Db          49 GCTTTCCAGGCCC--TAAAGAGGGCCCTAACCCAAGCCCCAGTGTTAAGCT 1
```

RESULT 131
DA673983
LOCUS DA673983 592 bp mRNA linear EST 06-NOV-2005
DEFINITION DA673983 NETRP2 Homo sapiens cDNA clone NETRP2004674 5', mRNA sequence.
ACCESSION DA673983
VERSION DA673983.1 GI:80920664
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

```
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NETRP2004674"
                     /cell_type="neutrophils"
                     /clone_lib="NETRP2"
                     /note="Vector: pME18SFL3; primary culture, neutrophils"

  Query Match             6.2%;  Score 466.3;  DB 1;  Length 592;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  509;  Conservative   48;  Mismatches   35;  Indels   21;  Gaps    4;

Qy       2840 CTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACC 2899
Db          1 CTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCACCTGGACTGTTTTACC 51

Qy       2900 CCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCA 2959
Db         52 CCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCA 111

Qy       2960 RTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYR 3019
Db        112 ATTGTCATAGCTGGACATTCTTGTTCTTCTGTATGGGGATGATTTAATTATAGCCACCCA 171

Qy       3020 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGG 3079
Db        172 TTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGG 231

Qy       3080 CTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT 3139
Db        232 C---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCT 282

Qy       3140 AARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTA 3199
Db        283 AAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTA 342

Qy       3200 TCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCG 3259
Db        343 TCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCG 402

Qy       3260 AAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAA 3319
Db        403 AATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAA 461

Qy       3320 CTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGG 3379
Db        462 CTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGG 521

Qy       3380 CCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTT 3439
Db        522 CCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTT 579

Qy       3440 TTSTTYATAYRTC 3452
Db        580 TTCTTTATATGTC 592


RESULT 132
DB381241/c
LOCUS       DB381241                508 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB381241 PLACE3 Homo sapiens cDNA clone PLACE3000261 3', mRNA
            sequence.
ACCESSION   DB381241
VERSION     DB381241.1  GI:83092303
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE3000261"
                     /tissue_type="placenta"
                     /clone_lib="PLACE3"
                     /note="Vector: pME18SFL3"

  Query Match             6.2%;  Score 465.8;  DB 1;  Length 508;
  Best Local Similarity   95.1%;  Pred. No. 0;
  Matches  485;  Conservative    8;  Mismatches   15;  Indels    2;  Gaps    2;

Qy       7072 TACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGC 7131
Db        508 TNCCGCANACCCTTGGACCGGCCTTGTAGCCCNCGATCTGATGTTAATGCCATCAAAGGC 449

Qy       7132 ACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGC 7191
Db        448 -CCCNTCTTGAGGAAATNTCAGCTGCACAACNTNTA-TACGCCCCAATTCAGCAGGAAGC 391

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
Db        390 AGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 331

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
Db        330 GGACTGAGAGNCAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCT 271

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
Db        270 GGGAAGGTGACCACATCCNCCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCA 211

Qy       7372 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 7431
Db        210 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 151

Qy       7432 ATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCG 7491
Db        150 ATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTNTTCG 91

Qy       7492 AGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 7551
```

```
Db         90 AGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 31

Qy       7552 CTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||||||||||||||||||||||||:||||
Db         30 CTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 133
DB378147/c
LOCUS       DB378147                478 bp    mRNA    linear   EST 04-DEC-2005
DEFINITION  DB378147 PLACE1 Homo sapiens cDNA clone PLACE1008489 3', mRNA
            sequence.
ACCESSION   DB378147
VERSION     DB378147.1  GI:83165921
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008489"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             6.2%;  Score 464.2;  DB 1;  Length 478;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  460;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       7104 ACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACC 7163
              |  ||  |||||||| || ||||||||||||||| ||||||||||||| |||||||||
Db        478 AANATNNGATGTTAANGANATCAAAGGCACCCCTCNTGAGGAAATCTCAGNTGCACAACC 419

Qy       7164 TCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAAC 7223
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        418 TCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAAC 359

Qy       7224 AGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTA 7283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTA 299

Qy       7284 GGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGG 7343
              |||||||:|||||||||:|||||||||:|||||||||||| ||||||||||||||||||
Db        298 GGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTACATCCACCTTTAAACACGG 239

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 7403
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        238 GGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 179

Qy       7404 AAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACA 7463
              ||||||||||||||||||||||||||||||:||||||:||||||||||||||||||||||
Db        178 AAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACA 119

Qy       7464 ATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCT 7523
              |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        118 ATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCT 59

Qy       7524 CCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          58 CCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 134
DB377997/c
LOCUS       DB377997                577 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB377997 PLACE1 Homo sapiens cDNA clone PLACE1008285 3', mRNA
            sequence.
ACCESSION   DB377997
VERSION     DB377997.1  GI:83148169
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008285"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             6.2%;  Score 463.5;  DB 1;  Length 577;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  525;  Conservative    7;  Mismatches   38;  Indels    5;  Gaps    5;

Qy       7009 TCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              | || ||| || | || |||:|||  | ||| || |||||||| ||||||||||| | |
Db        572 TTTTNCAGGATGGNAGGTGTAAAAAAANAAAGGGNGCCCAAGAGGCAGTCCAAGATT-AN 514

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              || ||||| |||||||||||||||||||| ||||||| ||| |||||||||| ||  |||||
```

```
Db        513 ATTTACCGNAGACCCCTGGACCGGCCTGGTAGCCCAAGATTTGATGTTAA-GANNTCAAA 455

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
              |||||||||||||| ||||||  | ||  |||| |||||  ||  || ||||||||||||||||||
Db        454 GGCACCCCTCCTGNGGAAATTTNAGGTGCANAACCTTTAATAAGCCCCAATTCAGCAGGA 395

Qy       7189 AGCAGTTAGAGCGGTSGTCGGCCAACC-TCCCCAACAGCACTTAGG-TTTTCCTGTTGAG 7246
              |||||||||||  |||  ||  |||||||| ||||||||||  |  |||||  ||||||||||||
Db        394 AGCAGTTAGAGNGGTTGTNGGCCAACCTTCCCCAACAGAAATTAGGTTTTTCCTGTTGAG 335

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
              |||||||||||||||||||||||||||||||||||||||||| |||| ||| |||||||||:||||||
Db        334 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCNTAGG-TGANTAAGAATCCCTAAGCC 276

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
              |||:||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        275 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCT 216

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 156

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGGATATAAACCCAAGT 7486
              ||||||| ||||||:|||||||||||||||||||||||||||||||| ||||||||||||||
Db        155 CAATCATNTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGGTATAAACCCAAGT 96

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 TTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 36

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||||||||||||||||||:||||
Db          35 TCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1
```

RESULT 135
DA864742

| | |
|---|---|
| LOCUS | DA864742 491 bp mRNA linear EST 13-NOV-2005 |
| DEFINITION | DA864742 PLACE7 Homo sapiens cDNA clone PLACE7018614 5', mRNA sequence. |
| ACCESSION | DA864742 |
| VERSION | DA864742.1 GI:82356370 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 491) |
| AUTHORS | Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. |
| TITLE | Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes |
| JOURNAL | Genome Res. 16 (1), 55-65 (2006) |
| PUBMED | 16344560 |
| COMMENT | Contact: Takao Isogai<br>FLJ Project (HRI Team)<br>Helix Research Institute<br>2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan<br>Tel: 81-438-52-3975<br>Fax: 81-438-52-3986<br>Email: flj-cdna@nifty.com<br>NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB. |

```
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE7018614"
                     /tissue_type="placenta"
                     /clone_lib="PLACE7"
                     /note="Vector: pME18SFL3"

  Query Match             6.2%;  Score 462.5;  DB 1;  Length 491;
  Best Local Similarity   91.9%;  Pred. No. 0;
  Matches  452;  Conservative   36;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       3821 ATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTC 3880
              |||||||||||||:|||||||||||||||||||||||||||||||||||:||||:||||
Db          1 ATAACTGTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTC 60

Qy       3881 CYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 3940
              |:||||||||||:|||| |||||||||||||| ||||||||||||||||||||||||||||
Db         61 CTTTGACTGATCCCGACC-TCAACTTGTATACTGGTGGAAGTTCCTTTGTAGAAAAAGGA 119

Qy       3941 CTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
              ||||||||||:|||||||||||||||||||||||||||||:|||||||||||||||||||
Db        120 CTTCGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTC 179

Qy       4001 ACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTA 4060
              ||||||||||||||||||:||||:|||||||||||||||:|||:|:|||||||||||||||
Db        180 ACTCCAGGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTA 239

Qy       4061 GGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
              |||||||:||||||||:|||||||||||||||||||:|:||||||:||||||| |||||
Db        240 GGAGAAGAAAAAAGGGCAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCAT 299

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMA 4180
              |||||||::|||||||||:|:||||||||||||||||||||||:|||:|||||||||||:|
Db        300 GCCCATGCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAA 359

Qy       4181 CATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTC 4240
              ||||||||||||||||||:|||||||:||||:|||||||||||||:|||||||||:||||
Db        360 CATCAGGAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTC 419

Qy       4241 TTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCA 4300
              |||||||||:|||||||||| |||||||||:||||||||||:|||:|||:|||||||||
Db        420 TTACACTGCCGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGAGAACTGCCAAGCA 479

Qy       4301 KATATTGAAGCM 4312
              :||||||||||:
Db        480 GATATTGAAGCC 491
```

RESULT 136
AZ517418/c

| | |
|---|---|
| LOCUS | AZ517418 521 bp DNA linear GSS 16-OCT-2000 |
| DEFINITION | RPCI-11-47O8.TV RPCI-11 Homo sapiens genomic clone RPCI-11-47O8, genomic survey sequence. |
| ACCESSION | AZ517418 |
| VERSION | AZ517418.1 GI:10826891 |
| KEYWORDS | GSS. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 521) |
| AUTHORS | Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C. |
| TITLE | BAC end sequences of library RPCI-11 |
| JOURNAL | Unpublished (1997) |
| COMMENT | Other_GSSs: RPCI11-47O8.TJ<br>Contact: Shaying Zhao<br>Department of Eukaryotic Genomics<br>The Institute for Genomic Research |

```
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            This BAC end was generated during the R&D process and may have
            higher chance of clone tracking errors.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .521
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7518007"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-47O8"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             6.2%;  Score 461.8;  DB 1;  Length 521;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  460;  Conservative   32;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       2709 GTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAA 2768
             || | | ||  || |||||||||| |||||||||| ||::|||||||||||||||||| |
Db        521 GTTCCTCTAACCCAAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGCA 462

Qy       2769 GCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCT 2828
             ||| ||||||||||:|||| |||||||||| ||| || ||||| ||||||||||||||||
Db        461 GCAAAGTGGTTTACAGTCCTAGACCTTAAGGGTGCTTTTTTCTGTATCCCTGTACATCCT 402

Qy       2829 GACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGG 2888
             |||||||| |||||||||||||||||||| ||| |||||||:|||||||| ||||||||
Db        401 GACTCTCACTTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACATCTCAAGTCACCTGG 342

Qy       2889 ACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
             |||:|||||||||||||||||| |||||||:|||||||||||||||||||||||||||||
Db        341 ACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAA 282

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||| ||||||||||:|||:| |||||||||| ||
Db        281 GACTTGAGCCAGTTCTCATACCTGGACAATCTTGTCCTTCAGTATGGGGATGATTTAATT 222

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
             ||:||: ||::|||||||||||||||||||||||||||||||:|||||:|||||||||
Db        221 TTGGCCACCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTC 162

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
             ||:|||||||| |||||:|||||||| :|:|:|||||:|||||||||||||||||||||
Db        161 GCTACCTGTGGTTACAAGGTTTCCAAATGAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 102

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
             ||||||||:||||:|||||||||||||||||:|||||||||||||||:::|||||||||
Db        101 TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGCATCCAGCCA 42

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAG 3229
             |||||||||||| ||||||:|||||||||||||||||||||
Db          41 ATACTGGCTTATTCTCATCCCAAAACCCTAAAGCAACTAAG 1


RESULT 137
DB380923/c
LOCUS       DB380923                533 bp    mRNA    linear   EST 11-DEC-2005
DEFINITION  DB380923 PLACE2 Homo sapiens cDNA clone PLACE2000161 3', mRNA
            sequence.
ACCESSION   DB380923
VERSION     DB380923.1  GI:83566989
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 533)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .533
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000161"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"

  Query Match             6.1%;  Score 456.3;  DB 1;  Length 533;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches  488;  Conservative    9;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTA--GCCCACGATC 7109
             | ||||||||| || | | | | | |||||| ||| ||||| || ||   ||      ||
Db        533 TCCAGTCCAAGNTTANGTTNTCCNGAAGACCCTTGGNCCGGCNTGGTAAGCCCACGATTC 474

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             |||||||||| |||||||||||||||| |    | ||       ||||||||||||||||
Db        473 TGATGTTAATGCCATCAAAGGCACCCCTTCNTGAGGAAATCTCAGTGCACAACCTCTACT 414

Qy       7170 AC-GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCAC 7228
              || |||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        413 ACGGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCAC 354

Qy       7229 TTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 7288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 TTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 294

Qy       7289 AYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 7348
             |:|||||||:|||||||:|||||||||||||||||||||||||||||||||||||||||
Db        293 ACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 234

Qy       7349 GCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGAC 7408
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGAC 174

Qy       7409 AGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGAT 7468
             |||||||||||||||||||||||:||||||:|||||||||||||||||||||||||||||
Db        173 AGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGAT 114
```

```
Qy       7469 CGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTT 7528
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        113 CGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTT 54

Qy       7529 TGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         53 TGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 138
DB381613/c
LOCUS       DB381613                507 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB381613 PLACE4 Homo sapiens cDNA clone PLACE4000496 3', mRNA
            sequence.
ACCESSION   DB381613
VERSION     DB381613.1  GI:83095127
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000496"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"

  Query Match             6.1%;  Score 454.2;  DB 1;  Length 507;
  Best Local Similarity   92.7%;  Pred. No. 0;
  Matches  470;  Conservative    8;  Mismatches   27;  Indels    2;  Gaps    1;

Qy       7077 CAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCC 7136
             || |||||||| |||||||  |||||| | || |||||||||| ||||||||||||||||
Db        507 CANACCCCTGGNCCGGCCTNGTAGCCCCCAATTTGATGTTAATGNCATCAAAGGCACCCC 448

Qy       7137 TCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTA 7196
             |||||||||||| |||| |||||||||| || ||||||||||||||||||||||||||||
Db        447 TCCTGAGGAAATNTCAGNTGCACAACCTTTANTACGCCCCAATTCAGCAGGAAGCAGTTA 388

Qy       7197 GAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC--TGTTGAGATGGGGGA 7254
             || ||||:|||||||||||||||||||  |||||| |||||   ||| |||||||||
Db        387 GANCGGTCGTCGGCCAACCTCCCCAACAGNCCTTAGGGTTTCCCTGGTTAAGATGGGGGC 328

Qy       7255 CTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGG 7314
                | | || |||| |||| |||||||| ||||||||:||||||||| |||||||||:||||

Db        327 TTAAAAGCCAGGNCTAGTTGGATTTCTTAGGCTGATTAAGAATCCNTAAGCCTAGCTGGG 268

Qy       7315 AAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC 7374
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        267 AAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATC 208

Qy       7375 AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATY 7434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        207 AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATC 148

Qy       7435 TATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGC 7494
             ||||||:||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        147 TATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGC 88

Qy       7495 CGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTA 7554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 CGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTA 28

Qy       7555 TTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||||||||||||||||||||:||||
Db         27 TTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 139
BX388917/c
LOCUS       BX388917                564 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX388917 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 3-PRIME, mRNA sequence.
ACCESSION   BX388917
VERSION     BX388917.2  GI:46845081
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30461045.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAB005ZC06_CS00428_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YJ18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             6.0%;  Score 449.8;  DB 1;  Length 564;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  486;  Conservative   10;  Mismatches   55;  Indels    2;  Gaps    1;

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             ||| || ||| || || | || ||||   | | |  |  | ||     | ||| ||   |
```

```
Db        564 CTGGGGAACCACAACCAAACGATGTTAAGAAGACAGTAAAAAACGAGAGAGCTACCAGAC 505

Qy       6453 ATTCTTAGTGCCCCCYAT--GRCCATCTACACTGAACAAGATTTATACAGTTATGTCATA 6510
              || || ||     ||:||  |:| |         || |||||||||  ||||||||||||
Db        504 ATACTAAGAAGAGCCTATGAGACAACAGCGCTGGAGCAAGATTTAATCAGTTATGTCATA 445

Qy       6511 TCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTA 6570
              | ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TTTAAGCCCCGCATCAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTA 385

Qy       6571 GGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTA 6630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTA 325

Qy       6631 TCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGAT 6690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 TCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGAT 265

Qy       6691 CAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACC 6750
              ||||||||||||||||||||||||||||:||||||||||||||||||||:|||||||
Db        264 CAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACC 205

Qy       6751 GCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 6810
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 145

Qy       6811 TCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGAR 6870
              ||||||||||||||||||||||||:|||||||||:||||||||||||||||:||||||:
Db        144 TCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAG 85

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||   |||
Db          84 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTNCNCCC 25

Qy       6931 TTCTTAGGACCTC 6943
                |  ||||||| |
Db          24 CTNCTAGGACCCC 12
```

RESULT 140
AQ541847/c
LOCUS AQ541847 625 bp DNA linear GSS 19-MAY-1999
DEFINITION RPCI-11-344D13.TV RPCI-11 Homo sapiens genomic clone RPCI-11-344D13, genomic survey sequence.
ACCESSION AQ541847
VERSION AQ541847.1 GI:4872303
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-344D13.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .625
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7631796"
/db_xref="taxon:9606"
/clone="RPCI-11-344D13"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

Query Match 6.0%; Score 448.7; DB 1; Length 625;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 559; Conservative 1; Mismatches 61; Indels 11; Gaps 7;

```
Qy       1867 TTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCC 1926
              |||||||||||| |||||||||||||||||| ||||||||||||||||||||||||| |
Db        625 TTGCTTCCAGTGTGGTCTACAAGGACACTTTANAAAAGATTGTCCAAGTAGAAGTAAACT 566

Qy       1927 GCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986
              ||||| || |||||||| ||||| |||||| ||||||||||||||||||||||||||||||
Db        565 GCCCCCTCATCCATGCCTCTTATGTCAAGGAAATCACTGGAAGGCCCACTGCCCCAGGGG 506

Qy       1987 ACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTG 2046
                 ||| || | ||||||| |||||||||||| ||||||||||||||||||||||||||||
Db        505 -TGAAGTTCCTCTGAGTCAAAAGCCACTAACCGGATGATCCAGCAGCAGGACTGAGGGTG 447

Qy       2047 CCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCAT 2106
              |||||||||| | ||| |||||||||||| |||| |||||||||| |||||||||| ||
Db        446 CCTGGGGCAACCACCAGCCCATGCCATCATCCTCGCAGAGCCCTGAGTATGCTTGACTAT 387

Qy       2107 TGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTT 2163
              |||||||| ||||| |   |||||||||||||||||| || ||||| ||||||||||
Db        386 TGAGGGCCGGGAAGTTAACTGTCTCCTGGACACTGGTGTGGCCTTCTCAGTCTTACTCTC 327

Qy       2164 CTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGC 2223
              |||||| |||||||||||||||| ||||| ||||||   | | |||  | || |||   |
Db        326 CTGTCCTGGACAACTGTCCTCCGGATCTATCACTA---TCTGAGGGGTCCTAGGACAATC 270

Qy       2224 AGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCA 2283
              | |||||||||| ||:||||||||||||||||| || ||| |||| ||| | ||||||||
Db        269 ACTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTAGGGAACTTCACTCTTTTCA 212

Qy       2284 CATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCA 2343
              |||||||||||||||||| |||||||||||| |||| ||||||||||||||||||||||
Db        211 CATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCA 152

Qy       2344 AAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCT 2403
              |||||||| |||||||||| ||||||||||||||||||||||||||||||||| ||| |
Db        151 AAAGCAGGAGCCATTATACATCTGAACATAGGAGAAGGAACACCCGTTTGTTGT-CCCTT 93

Qy       2404 GCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAA 2463
              |||||||||||||||||||| |||||| |||||||||||||||||| || ||| |||
Db         92 GCTTGAGGAAGGAATTAATCCTAAAGTCTAGGCAACAGAAGGACAATATAGATGAG-CAA 34

Qy       2464 AGAATGCCCGTCCTGTTCAAGTTAAACTAAAG 2495
              ||||| ||| | ||||||||||||||||||||
Db          33 AGAATACCCATTCTGTTCAAGTTAAACTAAAG 2
```

RESULT 141
AG195724/c
LOCUS AG195724 520 bp DNA linear GSS 06-MAR-2004
DEFINITION Pan troglodytes DNA, clone: RP43-074P10.T7, genomic survey sequence.
ACCESSION AG195724
VERSION AG195724.1 GI:45227900
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                Sequencing: T7
             LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-074P10.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"

  Query Match             5.9%;  Score 444.9;  DB 1;  Length 520;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  453;  Conservative   28;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
             ||||:||||||||||||||||||||||||||||:||||:: ||||||||||||||||||
Db        520 AGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGACATACTCACTGCTAAAG 461

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
             ||||||||||||||||||||:||||||||||:||||||||| | ||||||:| |
Db        460 AAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTAATGCTTGAAGGAC 401

Qy       3746 CAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATG 3805
             |||||||||:||||  |||||||||||||||||| ||||:| ||||||||||||||||||
Db        400 CAGTGCTGCGACTGCACACTTGTGCAACTCTTAAACCAGCCACATTTCTTCCAGACAATG 341

Qy       3806 AAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
             ||||||||||:||   ||||||||||||:||||| |||||| ||| |||||| || |||
Db        340 AAGAAAAGATAGAAAGGAACTGTCAACAAGTAATTGCTCAAAACTACGCCACTTGAAGGG 281

Qy       3866 ACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCC 3925
             |||||:||||:|||||:||||||||||:|||| |||||||||||| ||||||||||
Db        280 ACCTTTTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTAATGGAAGTTCC 222

Qy       3926 TTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTT 3985
             | |||||||| ||||||||||:||||||||||||||||||||||||||||||||:||
Db        221 TCTGTAGAAAACGGACTTCGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTT 162

Qy       3986 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYT 4045
             ||||||||||||||||||||||||||||||:||||:||| ||||||||||:|||:|
Db        161 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAATAGCCCTCACT 102

Qy       4046 KGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCT 4105
             ||||||||||||||||||||:|||| |||:||||||||||||||||||:|:||||||
Db        101 CGGGCACTAGAATTAGGAGAAGGAAAATGGGTAAATATATATACAGACTCTAAGTATGCT 42

Qy       4106 YACCTAGTCNTCCATGCCCATGMRGCAA 4133
```

```
             :|||||||| |||||||||  |::||||
Db         41 TACCTAGTCCTCCATGCCCGCGCAGCAA 14


RESULT 142
DA835505
LOCUS       DA835505                562 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DA835505 PLACE6 Homo sapiens cDNA clone PLACE6000784 5', mRNA
            sequence.
ACCESSION   DA835505
VERSION     DA835505.1  GI:83063421
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6000784"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

  Query Match             5.9%;  Score 443.5;  DB 1;  Length 562;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  448;  Conservative   10;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        105 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 164

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 224

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 284

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        285 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 343

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
```

```
                 ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        344 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 403

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
                 ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        404 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 463

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
                 ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        464 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 523

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCAC 5810
                 |||||||||||||||||||||||||||||||||||||
Db        524 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCAC 562


RESULT 143
BQ365363
LOCUS       BQ365363                572 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  MR2-GN0027-040900-001-e09 GN0027 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365363
VERSION     BQ365363.1  GI:21040875
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR2&t2=MR2-GN0027-
            040900-001-e09&t3=2000-09-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 59
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .572
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0027"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.9%;  Score 439.7;  DB 1;  Length 572;
  Best Local Similarity   93.2%;  Pred. No. 0;
  Matches  508;  Conservative   11;  Mismatches   19;  Indels    7;  Gaps    6;

Qy       6654 GGAACGGGTCGCCGACTCCC-TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAG 6712
                 || ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         28 GGTACGGGTCGCCGACTCCCGTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAG 87

Qy       6713 TAGTCCTTCRAAATCGAAGAGCTTTAG-ACTYGCTAACCGCTGARAGAGGGGGAACCTGT 6771
                 |||||||||: |||||||||||||||| |||:||||||||||||:||||||||||  ||
Db         88 TAGTCCTTCAGAATCGAAGAGCTTTAGAACTTGCTAACCGCTGAAAGAGGGGGAACTGGT 147

Qy       6772 TTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAA 6831
                 |||  |||||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db        148 TTA--TTTAGGGGAAGAATGCGGTTATTATGTTAATCAATCGGGAATCGTCACTGAGAAA 205

Qy       6832 GTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCC 6891
                 |||:||||||||:|||||||||||| |||:||||||:|||||||||||||| ||||||
Db        206 GTTAAAGAAATTCGAGATCGAATACAAGGTAGAGCAGAGGAGCTTCGAAACACGGGACCC 265

Qy       6892 TGGGGCCTCCTCAGCCRAT-GGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGC 6950
                 ||||||||||||||:|| |||||||| ||||||||||||||||||||||||||||||||
Db        266 GGGGGCCTCCTCAGCCAATGGGATGCCCGGGATTCTCCCCTTCTTAGGACCTCTAGCAGC 325

Qy       6951 TATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTC 7010
                 |||||||| |||||||||||||||||| |||||||:|||||| |||||||| |||||
Db        326 TATAATATGGCTACTCCTCTTTGGACCCGGTATCTTTAACCTCCTGGTTAACTTGGTCTC 385

Qy       7011 TTCCAGAATCGAAGCT-GTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGA 7069
                 |||||||||||||||| ||:|||||||||||||||||||||||||||||||||||||||
Db        386 TTCCAGAATCGAAGCTGGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGA 445

Qy       7070 TCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAG 7129
                 |||||||||||||||||| |||||||:||||||| |||||||||||| ||||||||||
Db        446 TCTACCGCAGACCCCTGGACTGGCCTGCTAGCCCAGGATCTGATGTTACTGACATCAAAG 505

Qy       7130 GCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAA 7189
                 |||||||| |||||||||||||||||||||||||||| |||||||| ||||||||||||
Db        506 GCACCCCTCTTGAGGAAATCTCAGCTGCACAACCTCTAATACGCCCC-ATTCAGCAGGAA 564

Qy       7190 GCAGT 7194
                 |||||
Db        565 GCAGT 569


RESULT 144
DB381029/c
LOCUS       DB381029                509 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB381029 PLACE2 Homo sapiens cDNA clone PLACE2000344 3', mRNA
            sequence.
ACCESSION   DB381029
VERSION     DB381029.1  GI:83090891
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
```

```
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .509
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000344"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"

  Query Match             5.9%;  Score 439.3;  DB 1;  Length 509;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  475;  Conservative    9;  Mismatches   23;  Indels    3;  Gaps    3;

Qy       7074 CCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCAC 7133
              ||||| |||| ||| ||||||| |||||| ||||| ||||| |||| ||||||||||||
Db        509 CCGCANACCCTTGGNCCGGCCTTGTAGCCC-CGATCNGATGTNAATGCCATCAAAGGCAC 451

Qy       7134 CCCTCC-TGAGGAAATCTCAGCTGCACAACCTC-TACTACGCCCCAATTCAGCAGGAAGC 7191
              |||| | |||| ||||||||| |||||||| |   || |||| ||| | |       |||||
Db        450 CCCTTCTTGAGNAAATCTCAGTTGCACAACNTNTTATTACGGCCCCAATTCAGCAGAAGC 391

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 AGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 331

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
              ||||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||:|
Db        330 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCT 271

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        270 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCA 211

Qy       7372 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 7431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 151

Qy       7432 ATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCG 7491
              ||:||||||:|||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        150 ATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCG 91

Qy       7492 AGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 7551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          90 AGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 31

Qy       7552 CTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||||||||||||||||||||||:||||
Db          30 CTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 145
AQ878055
LOCUS       AQ878055                506 bp    DNA     linear   GSS 09-NOV-1999
DEFINITION  HS_2160_A2_H04_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2160 Col=8 Row=O, genomic survey
            sequence.
ACCESSION   AQ878055
VERSION     AQ878055.1  GI:6309522
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2160  row: O  column: 8
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 506.
FEATURES             Location/Qualifiers
     source          1. .506
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2160 Col=8 Row=O"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.8%;  Score 437.6;  DB 1;  Length 506;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  446;  Conservative   36;  Mismatches   22;  Indels    2;  Gaps    2;

Qy       3426 ACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCT 3485
              ||:|||||||||| ||:||:|||::|||||| || ||||||:||:|||||: |||||||
Db          2 ACAGGGCAAGACTNTTCTTTATATGTCACAG-AAGAAACAGGAATAGCTCTATGAGTCCT 60

Qy       3486 TACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGT 3545
              |||||||:|| :|||||:|||||||||||::||||::||||||:||||||||:||||||
Db         61 TACACAGGTCTGAGGGATGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGT 120

Qy       3546 AGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATC 3605
              |||||||||||||||:|:||||||||||:|||||||||  ||||||||||::|||||||
Db        121 AGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATC 180

Qy       3606 TGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAA 3665
              ||||| |||||||||||||||||:|||||||||||||||||| ||||||||||:|||||
Db        181 TGAAGTAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGAACATCTCATGATGTGAA 240

Qy       3666 YRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCA 3725
              ::||||| |||||||||||||||||||||||||||||||||||||||:|||||||||||: ||
Db        241 TGGCATAGTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTACTTAAATACCA 300

Qy       3726 GGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGY 3785
              ||||||||||||||||:||||||||||||:|||||||||||||||||||||||||||||||
Db        301 GGCTCTATTACTTGAAAGGCCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGA 360

Qy       3786 CNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCA 3845
              | |||||||||||||||||||| |||||:||:||||||| ||||||:||||| ||||
Db        361 CACATTTCTTCCAGACAATGAAGAANAGATAGAACATAACTGCCAACAAGTAATTGCTCA 420

Qy       3846 AACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACT 3905
              |||||||||||||||| |||||||| ||||:|||||:||||||||||: ||| ||||||
Db        421 AACCTATGCCACTCGANGGGACCTTNTAGAGGTTCCCTTGACTGATCCCAACC-TCAACT 479

Qy       3906 TGTATACTGATGGAAGTTCCTTTGTA 3931
              |||||||||||||| | || ||||
Db        480 TGTATACTGATGGAANGTNCTCTGTA 505


RESULT 146
AQ553777
LOCUS       AQ553777                565 bp    DNA     linear   GSS 28-MAY-1999
DEFINITION  RPCI-11-414M9.TV RPCI-11 Homo sapiens genomic clone RPCI-11-414M9,
```

```
            genomic survey sequence.
ACCESSION   AQ553777
VERSION     AQ553777.1  GI:4912954
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7658888"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-414M9"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             5.8%;  Score 436.7;  DB 1;  Length 565;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  472;  Conservative   51;  Mismatches   40;  Indels    5;  Gaps    4;

Qy       3230 RGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAG 3289
              :|::|||||||||:|||:|||:|||||||  ||:|||||||  |||||||: || :||||||
Db             1 GGAATTCCTTGGCATAACAGGCTTCTGCCAAATATGGATT-CCCAGGTACGGTGAAATAG 59

Qy       3290 CCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGAT 3349
              |||||:|||||:| |||:| ||||||||||||| |||||  ||||||||||||:|||||
Db            60 CCAGGCCATTATACACACTTATTAAGGAAACTCAAAAAGCGAATACCCATTTAGTAAGAC 119

Qy       3350 GGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCC 3409
              ||:|:|||||:::|||  ||||||||||||   ||  | | |||  |||||||  :|||
Db           120 AGACACCTGAAGCAGAAGCAGCTTTCCAGGCCTTAAAGAAGGCCCT--AACCCAAACCCC 177

Qy       3410 AGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG-AAAAAAACAGRA 3468
              |||||||||:|||||||||:|||||||||||||:||:|||::||||| ||||||||||:|
Db           178 AGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAGAAAAAAAACAGGA 237

Qy       3469 AYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTG 3528
              |:|||||:|||||  |||||||  :|||:|||||:||||||||||::|||||::|||||
Db           238 ATAGCTCTAGGAGTTCTTACACAAGTCCAAGGGACGAGCTTGCAACCTGTGGCATACCTG 297

Qy       3529 ASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCA 3588
              |:|||||||:  |||||||| ||||||||:|:||||||||:  ||||||  |  |||
Db           298 AGTAAGGAAATAGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACAGGTAGTTGCTGCA 357

Qy       3589 GTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGG 3648
              | ||||||::|||||||||||||||||||||||||||||:||| ||||||||||||||
Db           358 GCAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAAATCTTACTGTGTGG 417

Qy       3649 ACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAAC 3708
              ||||||||||:|||||::|||| |||||| |||||||||||||||||||||||||||||
Db           418 ACATCTCATGATGTGAACGGCATGCTCACTACTAAAGGAGACTTGTGGCTGTCAGACAAC 477

Qy       3709 Y-GTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTG 3767
              :  |||||||||||:|||||||||||||||||||:||||||||  | :||||  ||||||
Db           478 CATTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGTGGTAACTGCACACTTG 537

Qy       3768 TGCAACTCTTAACCCAGYCNCATTTCTT 3795
              |||||||||||||||| :| ||||| ||
Db           538 TGCAACTCTTAACCCAACCACATTTTTT 565


RESULT 147
BE732673
LOCUS       BE732673                440 bp    mRNA    linear   EST 15-SEP-2000
DEFINITION  601571305F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925728 5',
            mRNA sequence.
ACCESSION   BE732673
VERSION     BE732673.1  GI:10146665
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM752  row: j  column: 01
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3925728"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_21"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

  Query Match             5.8%;  Score 436.4;  DB 1;  Length 440;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches  431;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 60

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 GTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 120

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
```

```
                |||||||||||||||||||||||||||:|||||||||:|||||||||:||||||||||||
Db        121 ACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGA 180

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGC 7381
                |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        181 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGC 240

Qy       7382 TCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCM 7441
                |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||:
Db        241 TCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCC 300

Qy       7442 TGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAAC 7501
                |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        301 TGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAAC 360

Qy       7502 GGCAACCCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACT 7561
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACT 420

Qy       7562 CTATTAAATCTTGCARCTGC 7581
                ||||||||||||||:||||
Db        421 CTATTAAATCTTGCAACTGC 440


RESULT 148
BX409035/c
LOCUS       BX409035                484 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX409035 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
            3-PRIME, mRNA sequence.
ACCESSION   BX409035
VERSION     BX409035.2  GI:46956493
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30766926.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAD005ZE05_AD00422_1&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .484
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE012YJ24"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

  Query Match             5.8%;  Score 433.2;  DB 1;  Length 484;
  Best Local Similarity   95.5%;  Pred. No. 0;
  Matches  462;  Conservative    3;  Mismatches   17;  Indels    2;  Gaps    2;

Qy       6261 GGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAAC 6320
                ||||| ||||||| | |||| ||||||||||||||||||||||||||||||||||||||
Db        484 GGAAAGAACCCATCCGTCAAGCCTCACCTGTGTAAAATTTAGCAATACTACATACACAAC 425

Qy       6321 CAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGG 6380
                |||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        424 CAACTCACAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCACAGA 365

Qy       6381 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTAT 6440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTAT 305

Qy       6441 GTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAG 6500
                |||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        304 GTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAA 245

Qy       6501 TTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGC 6560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGC 185

Qy       6561 AGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTA 6620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTA 125

Qy       6621 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACT-CCCTGGTCA 6679
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        124 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCCTGGTCA 65

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAG-CAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||||| ||||||||| ||:|||| |     |||
Db          64 CCTTGCAAGATCAACTTAACTCCCTAGCAGCCAGTAGTCCCTCAAAATNCGAAGAGCTTA 5

Qy       6739 GACT 6742
                ||||
Db           4 GACT 1


RESULT 149
AQ384981
LOCUS       AQ384981                596 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-138L12.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-138L12,
            genomic survey sequence.
ACCESSION   AQ384981
VERSION     AQ384981.1  GI:4356004
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-138L12.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
```

```
     source          1. .596
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7552883"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-138L12"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             5.8%;  Score 433.2;  DB 1;  Length 596;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  524;  Conservative   12;  Mismatches   57;  Indels   10;  Gaps    6;

Qy       1059 CCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTT 1118
             |||||||||||||||||| |:||||::|||| | |||||||||||||| |||| |||| |
Db          3 CCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCGGACTTAAAGCAAATTAAAATAGACCT 62

Qy       1119 AGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATT 1178
             |||||||| |||||||||:||||||||||||||||:||||||||| ||||||||||||||
Db         63 AGGTAAATCCTCAGATAACCCTGATGGCTATATTGATGTTTTACAGGGGTTAGGACAATC 122

Qy       1179 CTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATG 1238
             ||| ||||||| |||||||||||||| |||| |||||||||||||  ||||||||||||
Db        123 CTTCGATCTGATATGGAGAGATATA-ATGTTACTGCTAAATCAGACGTTAACCCCAAATG 181

Qy       1239 AGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGG 1298
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        182 AGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGG 241

Qy       1299 TCAATGGATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAG 1358
             ||||||     |||||| ||||| |||||||| |||||||||||||||||||||:||||
Db        242 TCAATG--ATAGGATGATAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAG 298

Qy       1359 TTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGC 1418
             |||||||| ||:||||||| ||||       | |         |    ||||||||||||||
Db        299 TTCCCAGTGTAGACCCTCACTGGG----ACGCAGAATCAGAACATGGAGATTGGTGCTGC 354

Qy       1419 AGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAA 1478
             |||||||||||||||||||||| :|||||||||||||||| |:|||| ||: ||| : ||
Db        355 AGACATTTGCTAACTTGTGTGCTGGAAGGACTAAGGAAAACCAGGAAG-AAGCCTACAAA 413

Qy       1479 TTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGG 1538
             ||  |||||||||||||| |||||||| |||| |||||||||||||||||| |||||||
Db        414 TTTTTCAATGATGTCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCTTTTCTGG 472

Qy       1539 AGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCC 1598
             ||||||||| |||||||||||||||| | ||||||||||||||||||  |||| |||
Db        473 AGAGACTAAAGGAGGCATTGAGGAAGCATACCTCTCTGTCACCTGACTCTATTGAAAGCC 532

Qy       1599 AACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCA 1658
             ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||| |
Db        533 AACTAATCTTAAAAGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTC 592

Qy       1659 AAA 1661
             |||
Db        593 AAA 595


RESULT 150
BX492778
LOCUS       BX492778               470 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp781G0323_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781G0323 5', mRNA sequence.
ACCESSION   BX492778
VERSION     BX492778.1  GI:32004591
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781G0323) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781G0323"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"

  Query Match             5.8%;  Score 432.4;  DB 1;  Length 470;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  421;  Conservative   30;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       3977 GGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATA 4036
             | |  |:|||||||||||||||||||||||||||||||||||:||||:||||||||||||
Db          2 GAATGACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTAGCAGAACTAATA 61

Qy       4037 GCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCT 4096
             |||:|||:|:|||||||||||| ||||||||:|  |||||:|||||||||||||||||||
Db         62 GCCCTCACTTGGGCACTAGAATGAGGAGAAGAAGGAAGGGCAAATATATATACAGACTCT 121

Qy       4097 RARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTA 4156
             :|:|||||:|||||||| |||||||||||::|||||||:|:||||||||||||||||||
Db         122 AAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAAAGAAAGGGAATTCCTA 181

Qy       4157 ACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTA 4216
             |||||:|||:||||||||||||:|||||||||||||||||||:|||||||:||||:|||
Db         182 ACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGGCTGTA 241

Qy       4217 CAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGG 4276
             ||||||||||:||||||||:||||||||||||:|||||||| |||||||||:|||||
Db         242 CAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAAGGAAAGGAAAGG 301

Qy       4277 GAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCT 4336
             ||||||:|||:|||:|||||||||:|||||||||:|||||||||||||||||||||||||
Db         302 GAAATAGAAGAGAACTGCCAAGCAGATATTGAAGCCAAAAGAGCTGCAAGGCAGGACCCT 361

Qy       4337 CCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAG 4396
             |||||||||||||||| ||||  ||| |||||||||||||||| ||||||||||||||
Db         362 CCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCCTCCGGGAAACCAAG 421

Qy       4397 CCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG 4440
             |||||||||||||||||||||||||||||||||||||||||||||
Db         422 CCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG 465


RESULT 151
CZ465902
LOCUS       CZ465902               560 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF757114TF Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_57114, genomic survey sequence.
```

```
ACCESSION   CZ465902
VERSION     CZ465902.1  GI:77949759
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 560)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .560
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_57114"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."

  Query Match             5.8%;  Score 431.2;  DB 1;  Length 560;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  481;  Conservative   32;  Mismatches   36;  Indels   16;  Gaps    4;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              |||||||| ||:|:||||||||||:|||||||||  ||||||||||::||||||||||
Db             1 GCAAAGGGCTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |||||||||||  ||||||||:|||||||||||||||||||||||||||| |:|||||::||
Db            61 GCAGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATCATGTGAATGGC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGC- 3728
              |||||||||||||||| |||||  || ||||||||||: ||| |||||||:||||||
Db           121 ATACTCACTGCTAAAGGGGACTTCGGGTTGTCAGACAACCATTTGCTTAAATATCAGGCC 180

Qy       3729 --TCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
                |||||||||||:||||||||||:||||  ||||||||||||||||     ||:|
Db           181 AGTCTATTACTTGAAGGGCCAGTGCTGCAACTGCACACTTGTGCAACTCTT-----AGCC 235

Qy       3787 NCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAA 3846
              |||||||||||||||||||| |||||||:| :|||||||||||||:| ||| |||||
Db           236 ACATTTCTTCCAGACAATGAAAAAAAGATAGAGCATAACTGTCAACAAGTGATTGCTCAA 295

Qy       3847 ACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGAC-------CT 3899
              |||||||||||| |||||||||:||||:|||||:|||||||||||: ||       |
Db           296 ACCTATGCCACTCAAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCCACGTCCCCACA 355

Qy       3900 TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATG 3959
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |:| ||||||
Db           356 TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAATGACTTCGAAAGGCGAGGTATG 415

Qy       3960 CAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTY 4019
              ||||||| ||||| |||||||||:|| | |||||||||||||||||| |||||| |||:
Db           416 CAGTGGTGAGTGACAATGGAATACTTCATAGTAATCCCCTCACTCCAGAAACTAGCGCTC 475

Qy       4020 AGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA-TTAGGAGAAGRAAAAAGGGYA 4078
              ||||:|||||||||||||:|||:|:|||||||||||    |||||||:|||||| |:|
Db           476 AGCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAACACAGGAGAAGGAAAAAGAGTA 535

Qy       4079 AATATATATACAGACTCTRARTATG 4103
              |||||||||||||| || |: :||||
Db           536 AATATATATACAGCCTATAGGTATG 560


RESULT 152
DA045520
LOCUS       DA045520                579 bp    mRNA    linear   EST 28-OCT-2005
DEFINITION  DA045520 BLADE2 Homo sapiens cDNA clone BLADE2008043 5', mRNA
            sequence.
ACCESSION   DA045520
VERSION     DA045520.1  GI:78255492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BLADE2008043"
                     /tissue_type="bladder"
                     /clone_lib="BLADE2"
                     /note="Vector: pME18SFL3"

  Query Match             5.7%;  Score 429.2;  DB 1;  Length 579;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  503;  Conservative   10;  Mismatches   64;  Indels    4;  Gaps    4;

Qy       5513 TCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATA 5572
              |||||||||| ||| |||||||||||| ||||||||| |||||||||||||| |||  ||
Db             1 TCTAGAGGATCTGCACCTGCTCTTCAAGCAACAACCATGAGGAAAGTAACTAGAATTGTA 60

Qy       5573 AATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCA 5632
              ||||||   | :|  || |      ||||||||||:||:||||:| |:|||| :|||||
Db            61 GATCCCCATGGCCCTCCCTTGTCATAATTTTTCTCTTTACTGTTCTCTTACCCCCTTTCA 120
```

```
Qy       5633 CTCTCACTGCACCCCCTCCATGCC-GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTT 5691
Db        121 CACTCACTGCACCTCCTCCATGCNTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTT 180

Qy       5692 CTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAG 5751
Db        181 CTATGAAGAATGCGGCTTCCCGAATATATTAATGCCCCATTGTATAGGAGTTTTTCTAAA 240

Qy       5752 GGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACT 5811
Db        241 TGAAACCCCACTTTCAC-CCCCACATCCATATGCCCCGCAACTGCTATAACTCTGCCACT 299

Qy       5812 CTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCT 5871
Db        300 CTTTGCATGCAAGCAAATACTCATTATTGGACAGGGAAAATAATTACTCCCAGTTGTCCT 359

Qy       5872 GGAGGACTTGGAGTCACTGTCTGTTGGA-CTTACTTCACCCAAACTGGTATGTCTGATGG 5930
Db        360 GG-GGGCTTGGAGCCACTGTCTGTTGGACCTTACTTCACCCATACCAGTATGTCTGATGG 418

Qy       5931 GGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCAC 5990
Db        419 GGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGGAAGTAATCTCCCAATTGAC 478

Qy       5991 CSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 6050
Db        479 TCAGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGACCTCTCAAAACTACATGAAAC 538

Qy       6051 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCA 6091
Db        539 CCTCCATACCCATACTCACCTGGTAAGCCTATTTAATACCA 579
```

```
RESULT 153
AA426511
LOCUS       AA426511                490 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  zw02e05.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:768128 5'
            similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AA426511
VERSION     AA426511.1  GI:2106774
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .490
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:768128"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-PacI;
                     Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid
                     DNA from three normalized libraries (melanocyte 2NbHM,
                     pregnant uterus NbHPU, and fetal heart NbHH19W) were
                     mixed, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from pools of 5,000 clones made from the same 3 libraries.
                     The pools consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
```

Query Match 5.7%; Score 427.8; DB 1; Length 490;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 465; Conservative 10; Mismatches 2; Indels 4; Gaps 4;

```
Qy       7019 TCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 7078
Db          1 TCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 60

Qy       7079 GACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 7138
Db         61 GACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 120

Qy       7139 CTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGA 7198
Db        121 GTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGA 180

Qy       7199 GCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGA 7258
Db        181 GCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGA 240

Qy       7259 GAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGG 7318
Db        241 GAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGG 300

Qy       7319 TGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAG 7378
Db        301 TGA-CACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAG 359

Qy       7379 AGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATT 7438
Db        360 AGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATT 419

Qy       7439 GCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGC 7498
Db        420 GCCTGAGA-CACAGCAGGAGGG-CAATGATCGGGATAT-AACCCAAGTCTTCGAGCCGGA 476

Qy       7499 A 7499
Db        477 A 477
```

```
RESULT 154
AA776439/c
LOCUS       AA776439                427 bp    mRNA    linear   EST 05-FEB-1998
DEFINITION  zj50h06.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:453755 3' similar to SW:ENV_SMRVH P21412 ENV
            POLYPROTEIN PRECURSOR ;, mRNA sequence.
ACCESSION   AA776439
VERSION     AA776439.1  GI:2835773
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
```

```
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 410.
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1390111"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:453755"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             5.7%;  Score 425;  DB 1;  Length 427;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches  422;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6402 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGT 6461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGT 368

Qy       6462 GCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCG 6521
             ||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCG 308

Qy       6522 CAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGG 6581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGG 248

Qy       6582 TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACT 6641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACT 188

Qy       6642 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 6701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 128

Qy       6702 CCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGG 6761
             ||||||||||||||||||||:||||||||||||||||||:||||||||||:|||||
Db        127 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGG 68

Qy       6762 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGT 6821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          67 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGT 8

Qy       6822 CACTGAG 6828
             |||||||
Db           7 CACTGAG 1


RESULT 155
DB380906/c
LOCUS       DB380906                498 bp    mRNA    linear   EST 11-DEC-2005
DEFINITION  DB380906 PLACE2 Homo sapiens cDNA clone PLACE2000129 3', mRNA
            sequence.
ACCESSION   DB380906
VERSION     DB380906.1  GI:83566955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000129"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"

  Query Match             5.6%;  Score 422.7;  DB 1;  Length 498;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  460;  Conservative    9;  Mismatches   24;  Indels    5;  Gaps    3;

Qy       7089 CCGGCCTGYTAGCCCACGATCTGATGTTAA---TGACATCAAAGGCACCCCTCCTGAGGA 7145
             |||||||| |   || |||| ||||  | |          || ||||||| || ||||||
Db        498 CCGGCCTGGTNAGCCCCGATNTGATNGTTAATGNCCATCAAAGGGCACCCNTCNTGAGGA 439

Qy       7146 AATCTC-AGCTGCACAACCTCTACTAC-GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
             | |||| |||||||||| | || ||| ||||||||||||||||||||||||||||||||
Db        438 ATTCTCAAGCTGCACAACTTTTANTACGGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 379

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
             :|||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        378 CGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTCGAGATGGGGGACTGAGAGAC 319

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||||||:|||||||||:|||||||||:|||||||||||||
Db        318 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC 259

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
             |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        258 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 199

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||:||
Db        198 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 139

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
             |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
```

```
Db        138 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGG 79

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 19

Qy       7564 ATTAAATCTTGCARCTGC 7581
              |||||||||||||:||||
Db         18 ATTAAATCTTGCAACTGC 1


RESULT 156
BQ024410/c
LOCUS       BQ024410                450 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-1-BB1p-auq-e-06-0-UI.s1 NCI_CGAP_Pl6 Homo sapiens cDNA clone
            UI-1-BB1p-auq-e-06-0-UI 3', mRNA sequence.
ACCESSION   BQ024410
VERSION     BQ024410.1  GI:19759689
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 14-226, >HERV17#LTR/Retroviral (matched compliment)
            214-256, >HERV17#LTR/Retroviral (matched compliment) 257-450,
            >LTR17#LTR/Retroviral (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .450
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BB1p-auq-e-06-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="Full Term"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Pl6"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl6 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human full term
                     TAG_LIB=UI-1-BB1p
                     TAG_SEQ=AGGAA"

  Query Match             5.6%;  Score 421.3;  DB 1;  Length 450;
  Best Local Similarity  96.8%;  Pred. No. 0;

  Matches  424;  Conservative   13;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        381 GTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCT 440
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        450 GTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCT 391

Qy        441 TGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGG 500
              ||||||||||||||:||| ||||||||||||||||||||:||||||||||||||||||||
Db        390 TGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGG 332

Qy        501 AAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAG 560
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        331 AAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAG 272

Qy        561 GACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATGGGAAACGTTCC 620
              ||||||||||:||||||||||:||||:::||||||||||:||||||||||||||||||||
Db        271 GACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATGGGAAACGTTCC 212

Qy        621 CCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCA 680
              ||||||||||||||||||||||||||||||||||||||:||||||||:||||||||||||
Db        211 CCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCA 152

Qy        681 GACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAA 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAA 92

Qy        741 GTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTG 800
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         91 GTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTG 32

Qy        801 AAGTGCCATAAGTACAAA 818
              ||||||||||||||||||
Db         31 AAGTGCCATAAGTACAAA 14


RESULT 157
B17809/c
LOCUS       B17809                  533 bp    DNA     linear   GSS 04-JUN-1998
DEFINITION  347L9.TVB CIT978SKA1 Homo sapiens genomic clone A-347L09, genomic
            survey sequence.
ACCESSION   B17809
VERSION     B17809.1  GI:2125558
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 533)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .533
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="A-347L09"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /clone_lib="CIT978SKA1"
```

```
                /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                CalTech Human BAC Library A1"

  Query Match             5.6%;  Score 420.6;  DB 1;  Length 533;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  442;  Conservative   35;  Mismatches   49;  Indels    6;  Gaps    1;

Qy       2728 TACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCC 2787
             ||| || |||||||||| ||::|||||||| |||||||| |||||||||||||| |:||
Db        533 TACTTAACCCTTATACTCTGCTTTCCCAAATGCCAGAGGAGGCAGAGTGGTTTATAGTCT 474

Qy       2788 TGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTG 2847
             ||||||||:|||||| ||| |||||||| ||||||||||||||||||||||||||||||
Db        473 TGGACCTTAAGGATGTCTTTTTCTGCATTCCTGTACATCCTGACTCTCAATTCTTGTTTG 414

Qy       2848 CCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGT 2907
             |||||||||||| |||| ||||||:  |||||||||||||||| |:|||||||||||||
Db        413 CCTTTGAAGATCCTTCTAACCCAATGTCTCAACTCACCTGGAATGTTTTACCCCAAGGGT 354

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCAT------TAGCCCAAGACTTGAGYCART 2961
             ||||||||||:|| |||||||| ||||||||||       ||||||| |||||||:||:|
Db        353 TCAGGGATAGCCCTCATCTATTCGGCCAGGCATCAACCCAAGCCCAACACTTGAGCCAGT 294

Qy       2962 YMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTT 3021
             ::|||||| || |||||| |||  || :|||  |||||||||||| | |:||:|||: |
Db        293 TCTCATACTTGAACACTCATGTTGTTTGGTACATGGATGATTTACATCTAGCCGCCCCTG 234

Qy       3022 CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCT 3081
             ||||| |||||||||||||||||||||||| :|| ||:|| |||||| |:|| |||||||
Db        233 CAGAAGCCTTGTGCCATCAAGCCACCCAAGTGCTTTTAAACTTCCTCACCACTTGTGGCT 174

Qy       3082 ACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAA 3141
             |||:|||||||||||:|:|: ||||:|||||| | |||||||||||||||||: |||
Db        173 ACAAGGTTTCCAAACCAAAGTCTCAGCTCTGCTTATAGCAGGTTAAATACTTAGGGTTAA 114

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
             :|||||||||||||||:||||||||| ||||||:::||||| ||||||||||||||||
Db        113 AATTATCCAAAGGCACCAGGGCCCTCAGCGAGGAATGTATCCATCCTATACTGGCTTATC 54

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
             || |||:||||||||| |||||| ||||:|: |||| ||||:|||:| |:||
Db         53 CTGATCCCAAAACCCTGAAGCAATTAAGAGGTTTCCATGGCATAACAAGCTT 2


RESULT 158
DN914892/c
LOCUS       DN914892                460 bp    mRNA    linear   EST 20-OCT-2005
DEFINITION  MCF7RNAL19F11TF Human MCF7 breast cancer cell line near full length
            normalized library (MCF7_EST) Homo sapiens cDNA clone
            MCF7_RNA_L_19_F11, mRNA sequence.
ACCESSION   DN914892
VERSION     DN914892.1  GI:77940349
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7067
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            Seq primer: T7 5'-GTAATACGACTCACTATAGGGC-3'.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_RNA_L_19_F11"
                     /sex="female"
                     /cell_line="MCF7"
                     /clone_lib="Human MCF7 breast cancer cell line near full
                     length normalized library (MCF7_EST)"
                     /note="Vector: pAL17; This library was constructed from
                     MCF7 breast cancer cell line by Evrogen JSC.
                     (http://www.evrogen.com) using their standard procedure.
                     See http://shark.ucsf.edu/~stas/report_MCF7.pdf for
                     details."

  Query Match             5.6%;  Score 420;  DB 1;  Length 460;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  411;  Conservative   32;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2778 TTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAA 2837
             ||||||:|||||| ||||:|||||||||| |||||||||||||||||||||||||||||||
Db        460 TTTACAGTCCTGGGCCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAA 401

Qy       2838 TTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTA 2897
             |||||||||||||||||||  |||||||||| :|||||||||||||||||||||:|||||
Db        400 TTCTTGTTTGCCTTTGAAGAACCTTCAAACCCCACATCTCAACTCACCTGGACTGTTTTA 341

Qy       2898 CCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGY 2957
             ||  |||||||||||||||:|||||||||||||||||| |||||||||| |||||||:
Db        340 CCGAAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAAACTTGAGC 281

Qy       2958 CARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCC 3017
             ||:|::|||||||||||||||||||||||:||| |||||||||||||||||:||:|||
Db        280 CAATTCTCATACCTGGACACTCTTGTCCTTCGGTACGTGGATGATTTACTTTTAGCTGCC 221

Qy       3018 YRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGT 3077
             ::|||||||||||||||||||||||||||||||||:|||||:||||||||| |:||||||
Db        220 CGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCACTACCTGT 161

Qy       3078 GGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGR 3137
             |||||||:|||||||||||:|:|:|||||:|||||||||||| ||||||  ||||||||:
Db        160 GGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGGAGGTTACGTACTTAGGG 101

Qy       3138 CTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCT 3197
             ||||:|||||||||||||||:|||||||||||||||:::|||||||||||||||||||
Db        100 CTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCT 41

Qy       3198 TATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCC 3237
             ||  |||||: |||||||||||||||||||:|::||||
Db         40 TAATCTCATCCTAAAACCCTAAAGCAACTAAGAGGGTTCC 1


RESULT 159
DN914893
LOCUS       DN914893                460 bp    mRNA    linear   EST 20-OCT-2005
DEFINITION  MCF7RNAL19F11TR Human MCF7 breast cancer cell line near full length
            normalized library (MCF7_EST) Homo sapiens cDNA clone
            MCF7_RNA_L_19_F11, mRNA sequence.
ACCESSION   DN914893
VERSION     DN914893.1  GI:77940351
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
```

```
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7067
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            Seq primer: SP6 5'-ATTTAGGTGACACTATAGAATAC-3'.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_RNA_L_19_F11"
                     /sex="female"
                     /cell_line="MCF7"
                     /clone_lib="Human MCF7 breast cancer cell line near full
                     length normalized library (MCF7_EST)"
                     /note="Vector: pAL17; This library was constructed from
                     MCF7 breast cancer cell line by Evrogen JSC.
                     (http://www.evrogen.com) using their standard procedure.
                     See http://shark.ucsf.edu/~stas/report_MCF7.pdf for
                     details."

  Query Match             5.6%;  Score 420;  DB 1;  Length 460;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  411;  Conservative   32;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2778 TTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAA 2837
              ||||||:|||||| ||||:|||||||||| |||||||||||||||||||||||||||||
Db          1 TTTACAGTCCTGGGCCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAA 60

Qy       2838 TTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTA 2897
              |||||||||||||||||||  |||||||||| :|||||||||||||||||||:|||||
Db         61 TTCTTGTTTGCCTTTGAAGAACCTTCAAACCCCACATCTCAACTCACCTGGACTGTTTTA 120

Qy       2898 CCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGY 2957
              ||  ||||||||||||||:|||||||||||||||||| |||||||||| |||||||:
Db        121 CCGAAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAAACTTGAGC 180

Qy       2958 CARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCC 3017
              ||:|::|||||||||||||||||||||||:||| |||||||||||||||||:||:|||
Db        181 CAATTCTCATACCTGGACACTCTTGTCCTTCGGTACGTGGATGATTTACTTTTAGCTGCC 240

Qy       3018 YRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGT 3077
              ::|||||||||||||||||||||||||||||||:|||||:||||||||| |:||||||
Db        241 CGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCACTACCTGT 300

Qy       3078 GGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGR 3137
              |||||||:|||||||||||:|:|:|||||:|||||||||||| ||||||  ||||||||:
Db        301 GGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGGAGGTTACGTACTTAGGG 360

Qy       3138 CTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCT 3197
              ||||:|||||||||||||||:|||||||||||:::|||||||||||||||||||||||
Db        361 CTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCT 420

Qy       3198 TATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCC 3237
              ||  ||||||: |||||||||||||||||||:|::||||
Db        421 TAATCTCATCCTAAAACCCTAAAGCAACTAAGAGGGTTCC 460


RESULT 160
AW971553/c
LOCUS       AW971553                522 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST383642 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW971553
VERSION     AW971553.1  GI:8161399
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 292
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGL"
                     /note="Vector: pBluescriptSKm"

  Query Match             5.6%;  Score 418.6;  DB 1;  Length 522;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  463;  Conservative    9;  Mismatches   48;  Indels    2;  Gaps    2;

Qy       7060 AAGACTAAGATCTACCGCAGACCCCT-GGACCGGCCTGYTAGCCCACGATCTGATGTTAA 7118
              | |  |||||||| |||| |||||||   || ||| ||:||||||| | |||| ||||||
Db        522 AGGTTTAAGATCTTCCGCGGACCCCTAAAACTGGCTTGCTAGCCCATGTTCTGGTGTTAA 463

Qy       7119 TGACATCAAAGGCACCCCTCCTGAGGAAA-TCTCAGCTGCACAACCTCTACTACGCCCCA 7177
              ||||||| ||| ||| |||||||||||| ||||| ||||||||||  || || ||||||
Db        462 TGACATCGAAGTCACTCCTCCTGAGGAAATTCTCAACTGCACAACCCTTATTATGCCCCA 403

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
              |||||||||||||||||||| ||| || ||||||||| |||||||||| ||||||
Db        402 GTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGGGTTTT 343

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
              ||||||||||  |||||||||||||||||||||||||||||||||||| ||:|||||||
Db        342 CCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTAAGAAT 283

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
              ||:||| |||||:|||||||||||| |||||||||||||||||||||||||||||||||
Db        282 CCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAACTTAG 223

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
              :|||||||  ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        222 CTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 163

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
              |||||||||||||||:|||||:|||||||||||  |||||||||| ||| |  |||||
Db        162 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGTGGGAGGGACAAGGATTGCAATATA 103

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
              |||||| |  |||||||| |||||||||| |||||||||| ||||||||||||||||||
Db        102 AACCCAGGCATTCGAGCCAGCAACGGCAACCGCCTTTGGGTCCCTTCCCTTTGTATGGGA 43

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCT 7579
              ||||||||||||  ||||||||||||||||||||||||||:||
Db          42 GCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACT 1


RESULT 161
DB322339/c
LOCUS       DB322339                516 bp    mRNA    linear   EST 04-DEC-2005
DEFINITION  DB322339 NETRP2 Homo sapiens cDNA clone NETRP2004674 3', mRNA
```

```
            sequence.
ACCESSION   DB322339
VERSION     DB322339.1  GI:83258662
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NETRP2004674"
                     /cell_type="neutrophils"
                     /clone_lib="NETRP2"
                     /note="Vector: pME18SFL3; primary culture, neutrophils"

  Query Match             5.5%;  Score 414.1;  DB 1;  Length 516;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches  459;  Conservative    9;  Mismatches   48;  Indels    3;  Gaps    2;

Qy       7063 ACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
             ||||||||| || |||||||||| || |  ||| :||||||| | ||  ||||||||||
Db        516 ACTAAGATCCACTGCAGACCCCT-GATCAACCTCCTAGCCCATGCTCCAATGTTAATGAC 458

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
             ||  |||||||||||||| |||||||||||| |||||||||| |||||| | |||||||||
Db        457 ATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGACCCAATTCA 398

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
             |||||||||||||||||| ||: || |||||||||||||||||||||| ||||||||||
Db        397 GCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGT 338

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
             ||||| |||||||||||||||||||||||||||||||||||||||||| ||:||||||||:||
Db        337 TGAGAGGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTA 278

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||:|||||||||||||| |||| |||||||||| |||||||| |||||:||||
Db        277 AGCCTAGCTGGGAAGGTGACCGCATCTACCTTTAAACATGGGGCTTGCCACTTAGCTCAC 218

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
             |||  |||||||||||||||||||| | |||||||||||||| | ||||||||||||||
Db        217 ACCCAACCAATCAGAGAGCTCACTAAGAAGCTAATTAGGCAAAAATAGGAGGTAAAGAAA 158

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
```

```
             ||||||||||: |||||:|||||||||||  |||||||||| |||| ||||||||||||
Db        157 TAGCCAATCATCAATTGCCTGAGAGCACAGTGGGAGGGACAAGGATCCGGATATAAACCC 98

Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
             | |  || ||| | |||| ||||| ||||||||||||||||||  ||||||||||||||||
Db         97 AGGCATTTGAGACAGCAATGGCAATCCCCTTTGGGTCCCCTCC--TTGTATGGGAGCTCT 40

Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||  |||||||||||||||||||||||||:||||
Db         39 GTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGC 1
```

```
RESULT 162
DB377653/c
LOCUS       DB377653                514 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB377653 PLACE1 Homo sapiens cDNA clone PLACE1007839 3', mRNA
            sequence.
ACCESSION   DB377653
VERSION     DB377653.1  GI:83145805
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 514)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1007839"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             5.5%;  Score 413.9;  DB 1;  Length 514;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  459;  Conservative    7;  Mismatches   45;  Indels    5;  Gaps    3;

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
             || ||||| |||||| ||||| |||    |   |||  ||  |||||||| || || ||
Db         514 ATTTACCGNAGACCCNTGGACGGGC--TGTTAGCCAAAATTNGATGTTAANGAAATNAAG 457

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTA--CGCCCCAATTCAGCAG 7186
             || |||| ||  | |||||| | || |||||||| | || ||    ||||||||||||||
Db         456 GGAACCCTTCNGGNGGAAATTTTAGGTGCACAACTTTTAATAACGCCCCCAATTCAGCAG 397

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT-CCTGTTGA 7245
             |||||||||||||| | || ||||||| |||||||||| | ||||||||| ||||||||
Db         396 GAAGCAGTTAGAGCGTTTGTNGGCCAACTTCCCCAACAGNANTTAGGTTTTCCCTGTTGA 337
```

```
Qy       7246 GATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGC 7305
             ||||||||| |||||||||||||||||||||||||| |||||||| |||||||||:|||||
Db        336 GATGGGGGANTGAGAGACAGGACTAGCTGGATTTCNTAGGCTGANTAAGAATCCCTAAGC 277

Qy       7306 CTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACC 7365
             ||||:|||||||||||||||||||||||||||||| ||||||||||||||||:|||||||
Db        276 CTAGCTGGGAAGGTGACCACATCCACCTTTAAACANGGGGCTTGCAACTTAGCTCACACC 217

Qy       7366 TGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAG 7425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAG 157

Qy       7426 CCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAG 7485
             |||||||||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        156 CCAATCATTTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAG 97

Qy       7486 TYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTT 7545
             |:||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 TTTTCGAACCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTT 37

Qy       7546 TTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||||||||||||| | |||||||| |||||:| ||
Db          36 TTCATGCTATTTCANTNTATTAAATNTTGCAACCGC 1
```

RESULT 163
AQ042339
LOCUS AQ042339 539 bp DNA linear GSS 14-JUL-1998
DEFINITION CIT-HSP-2326B23.TF CIT-HSP Homo sapiens genomic clone 2326B23, genomic survey sequence.
ACCESSION AQ042339
VERSION AQ042339.1 GI:3311533
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2326B23.TR °
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2326B23"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"

Query Match 5.5%; Score 413.8; DB 1; Length 539;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 424; Conservative 38; Mismatches 21; Indels 2; Gaps 2;

```
Qy       3416 AAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTC 3475
             |||:||||| | :|||||||||||||:||:|||::||| | |||||| |||:||:||||||
Db         51 AAGCTTGCCGATGGGGCAAGACTTTTCTTTATATGTCATA-AAAAAACCAGGAATAGCTC 109

Qy       3476 TRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGG 3535
             |:|||||||||||||||:|||:|||||: |||||||||||::||||::|||||| |||||
Db        110 TAGGAGTCCTTACACAGGTCCGAGGGACTAGCTTGCAACCCGTGGCATACCTGAATAAGG 169

Qy       3536 AAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
             |||:||||||||||||||||||:|:||||||||||:|||||||| ||||||||||
Db        170 AAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAG 229

Qy       3596 TYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTC 3655
             |::|||||||||||||||||| ||||||||||||:|||||| ||||||||||||||||||
Db        230 TCTTAGTATCTGAAGCAGTTTAAATAATACAGGGAAGAGATTTTACTGTGTGGACATCTC 289

Qy       3656 ATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTAC 3715
             ||||:||||||::|||||||| ||||||||||||||||||||||||||||||:||||||
Db        290 ATGATGTGAACGGCATACTGACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTAC 349

Qy       3716 TTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTC 3775
             ||||||:|||||||||||||||||:||||||||||||:|||| ||| | |||||||||
Db        350 TTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGCGCATTCGTGCAACTC 409

Qy       3776 TTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAAR 3835
             |||||||||:| |||||||||| ||||||||||||||||||:||:|||||||||||||||:
Db        410 TTAACCCAGTCCCATTTCTTCCGGACAATGAAGAAAAGATAGAACATAACTGTCAACAAG 469

Qy       3836 TAATTTCTCAAACCTATGCCACTCGA-GGGGACCTTYTAGARGTTCCYTTGACTGATCCY 3894
             ||||| |||||||||| || ||||| ||||||| |:||||:|||||:||||||||||:
Db        470 TAATTGCTCAAACCTACACCGCTCGAGGGGGACCCTTTAGAGGTTCCCTTGACTGATCCT 529

Qy       3895 GACCT 3899
             |||||
Db        530 GACCT 534
```

RESULT 164
AQ568446/c
LOCUS AQ568446 568 bp DNA linear GSS 01-JUN-1999
DEFINITION HS_5302_B1_C10_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=878 Col=19 Row=F, genomic survey sequence.
ACCESSION AQ568446
VERSION AQ568446.1 GI:4961666
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 568)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

```
            Plate: 878  row: F  column: 19
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 568.
FEATURES              Location/Qualifiers
     source           1. .568
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=878 Col=19 Row=F"
                      /sex="male"
                      /clone_lib="RPCI-11 Human Male BAC Library"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      Male blood DNA was isolated from one randomly chosen donor
                      and partially digested with a combination of EcoRI and
                      EcoRI Methylase. Size selected DNA was cloned into the
                      pBACe3.6 vector at EcoRI sites"

  Query Match             5.5%;  Score 413.6;  DB 1;  Length 568;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  442;  Conservative   29;  Mismatches    7;  Indels    8;  Gaps    4;

Qy       3648 GACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC-AGACA 3706
             ||||||||||||:|||| ::|||||||||||||||||||||||||||||||||| |||||
Db        528 GACATCTCATGAGGTGANCAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAAGACA 469

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
              ||:||||||||||||:||||||||||||||||:||||||||||||:||||||||||||
Db        468 ACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCAGTGCTGCAACTGTGCACTT 409

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
              ||||||||||||||||||:| ||||||||||||||||||     ||||||:||:||||||
Db        408 GTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATG-----AAGATAGAATATAACT 354

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
              ||||||||:||||||||||||||||||||||||||||||||||||:||||:|||||:||||
Db        353 GTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGA 294

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
              |||||||:|||| ||||||||||||||||||||||||||||||||||||||||| || |
Db        293 CTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACCTCAA 235

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
              ||||:||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        234 AAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGAAAGTAATCCCCTCACTCCA 175

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
              ||||||||||||:||||:|||||||||||||||:|||:|:||||||||||||||||||||
Db        174 GGAACTAGTGCTTAGCTGGCAGAACTAATAGCCCTCATTGGGGCACTAGAATTAGGAGAA 115

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
              |: |||||||:|||||||||||||||||||:|:||||||:||  |||| ||||||||||||
Db        114 GG-AAAAGGGTAAATATATATACAGACTCTGAGTATGCTCACTNAGTCCTCCATGCCCAT 56

Qy       4127 GMRGCA 4132
              |::|||
Db          55 GAGGCA 50


RESULT 165
AI598135/c
LOCUS       AI598135                490 bp    mRNA    linear   EST 12-MAY-1999
DEFINITION  tn14a10.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2167578 3'
            similar to contains PTR5.t1 PTR5 repetitive element ;, mRNA
            sequence.
ACCESSION   AI598135
VERSION     AI598135.1  GI:4607183
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1407   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .490
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2167578"
                      /tissue_type="anaplastic oligodendroglioma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Brn25"
                      /note="Organ: brain; Vector: pT7T3D-PacI; Site_1: Not I;
                      Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                      oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      Library is normalized, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."

  Query Match             5.5%;  Score 412.8;  DB 1;  Length 490;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  441;  Conservative   10;  Mismatches   33;  Indels   10;  Gaps    1;

Qy       7089 CCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAAT 7148
              || |||||:||||||  | || ||||||||||||||||||||||||||||||||||||||
Db         490 CCAGCCTGCTAGCCCTTGTTCCGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAAT 431

Qy       7149 CTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCG 7208
              |||| |||||||||| |||| | |||||| |||||||||||||||| |||||||||:|||
Db         430 CTCANCTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCA 371

Qy       7209 GCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGAC 7268
              | |||||||||||  |||||||| |||||||||||||||| |||| ||||||||||||||
Db         370 GTCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGAC 311

Qy       7269 TAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATC 7328
              |||||||||||||||||  ||:|||||||:|||||||:|||||||||||||| | ||
Db         310 TAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTC 251

Qy       7329 CACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAA 7388
              |||||||||||||||||||||||||||||:||||||| ||||||||||||||||||||||
Db         250 CACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAA 191

Qy       7389 AATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGC 7448
              |||||||||||  ||||| |||||||||||||||||||||| ||||:||| ||:|||||||
Db         190 AATGCTAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGC 131

Qy       7449 ACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACC 7508
              |||||| |||||||||||||||||||||||||| |  |||||||| |||||||| |||
Db         130 ACAGCAAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACC 71

Qy       7509 CCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAA 7568
```

```
                 | |||||||||||||||||||||||||||||| ||||          ||||||||||||||
Db         70 CTCTTTGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTCACTCTATTAA 21

Qy       7569 ATCTTGCARCTGCR 7582
                ||||||||:||||:
Db         20 ATCTTGCAACTGCA 7


RESULT 166
DB375490/c
LOCUS       DB375490                439 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB375490 PLACE1 Homo sapiens cDNA clone PLACE1004890 3', mRNA
            sequence.
ACCESSION   DB375490
VERSION     DB375490.1  GI:83098022
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004890"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             5.5%;  Score 409.5;  DB 1;  Length 439;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  417;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              | ||||| |  |     |||| || ||||||||||||||||||||||||||||||||||||
Db        437 GNAATCTNANGTNGNACACCTNTANTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 378

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              :||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 CGTCGGCC-ACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 319

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              |||||||||||||||||||||||:|||||||||:|||||||||:|||||||||||||||
Db        318 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC 259

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        258 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 199

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||:||
Db        198 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 139

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        138 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGG 79

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 19

Qy       7564 ATTAAATCTTGCARCTGC 7581
              |||||||||||||:||||
Db          18 ATTAAATCTTGCAACTGC 1


RESULT 167
AQ270891/c
LOCUS       AQ270891                537 bp    DNA     linear   GSS 03-NOV-1998
DEFINITION  HS_2047_A1_H02_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2047 Col=3 Row=O, genomic survey
            sequence.
ACCESSION   AQ270891
VERSION     AQ270891.1  GI:3823431
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2047  row: O  column: 3
            Class: BAC ends
            High quality sequence stop: 537.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2047 Col=3 Row=O"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.5%;  Score 408.9;  DB 1;  Length 537;
  Best Local Similarity   83.2%;  Pred. No. 0;
  Matches  422;  Conservative   44;  Mismatches   38;  Indels    3;  Gaps    2;

Qy       2916 AGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGAC 2975
              ||:|||||||||||||||||||||||| ||||||||||:||:|::|| |||||||||
Db        536 AGCCCCCATCTATTTGGCCAGGCATTAGCNCAAGACTTGAGCCAATTCTCCTACCTGGAC 477

Qy       2976 ACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGC 3035
              |||||||||||||:|||  |||||||||||||| ||:||: ||::|||||||||||||||
Db        476 ACTCTTGTCCTTCGGTACATGGATGATTTACTNTTAGCCACCCATTCAGAAACCTTGTGC 417
```

```
Qy       3036 CATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAA 3095
             ||||||||||||||||:|||||:| || ||| ||:|||||||||||||:||||||||||
Db        416 CATCAAGCCACCCAAGCACTCTTAATTTCCCTTGCTACCTGTGGCTACAAGGTTTCCAAA 357

Qy       3096 CSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGC 3155
             |:|:|:|||||:|||||||||||||||||||||||||||||:|||  :|||||||||||||
Db        356 CCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTANAATTATCCAAAGGC 297

Qy       3156 ACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACC 3215
             ||||:  |||||||||||||:::|||||||||||||||| ||||||||||:|||||||
Db        296 ACCAGAACCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACC 237

Qy       3216 CTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAG 3275
             |||||||||||||  :|  :|||||||||:|||:|||:|||||||  ||:|||||||  | |||
Db        236 CTAAAGCAACTAACAGCGTTCCTTGGCATAACAGGTTTCTGCCAAATATGGATT-CNCAG 178

Qy       3276 GTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATAC 3335
             ||:  ||:||||||||||| :|||||:|||||: |||||||||||||| | |||||||||
Db        177 GTACAGCAAAATAGCCAGACCATTAAATACACGAATTAAGGAAACTCANANAGCCAATAC 118

Qy       3336 CCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCT 3395
             |||||||:|||||||||:|:  |||||:::|||||||| |||||||||   |  | |  ||
Db        117 CCATTTAGTAAGATGGACACATGAAGCAGAAGTGGCTNTCCAGGCCCTANAGAAGGCGCT 58

Qy       3396 TAAACCCAAGYCCCAGTGTTAAGYTTG 3422
                ||| ||  |:||||||||||  :| |
Db         57 --AACACATGCCCCAGTGTTAGCTTAG 33
```

```
RESULT 168
BF993783/c
LOCUS       BF993783                414 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-011100-205-f08 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF993783
VERSION     BF993783.1  GI:12400106
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            011100-205-f08&t3=2000-11-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 376.
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
```

```
  Query Match             5.4%;  Score 407.2;  DB 1;  Length 414;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  406;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4770 AGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGC 4829
             ||||||||||:||||||||||||||||||:||||||||||||||||||||||||||||||
Db        414 AGTCTGGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGC 355

Qy       4830 CCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGG 4889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGG 295

Qy       4890 CTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGC 4949
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||
Db        294 CTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACNCAGGGAGTATCNCAGGC 235

Qy       4950 GTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGA 5009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 GTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGA 175

Qy       5010 GAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCAC 5069
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GAAAATGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCAC 115

Qy       5070 ATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGG 5129
             ||||| ||:|||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        114 ATGGCTTGCTCTGTTGCCTATAGCCTTAGAAAGAATCTGCAACTTTCCCCAAAAAGCAGG 55

Qy       5130 ACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGT 5183
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGT 1
```

```
RESULT 169
AI128526/c
LOCUS       AI128526                436 bp    mRNA    linear   EST 27-OCT-1998
DEFINITION  qc61h10.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1714147 3' similar to contains PTR7.b1 PTR5 repetitive
            element ;, mRNA sequence.
ACCESSION   AI128526
VERSION     AI128526.1  GI:3597040
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 720   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 428.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

/clone="IMAGE:1714147"
/dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGATTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

Query Match 5.4%; Score 406.8; DB 1; Length 436;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 402; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

```
Qy       7169 TACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCAC 7228
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        436 TACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCAC 377

Qy       7229 TTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 7288
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        376 TTAGGTTTTCCTGTTGAGATGGGGGACTGAAAGACAGGACTAGCTGGATTTCCTAGGCTG 317

Qy       7289 AYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 7348
             |:|||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||||
Db        316 ACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 257

Qy       7349 GCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGAC 7408
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGAC 197

Qy       7409 AGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGAT 7468
             |||||||||||||||||||||||||:|||||||:|||||||||||||||||||||||||||
Db        196 AGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGAT 137

Qy       7469 CGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTT 7528
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        136 CGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTT 77

Qy       7529 TGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||||||||||||| ||||||||||||||:||||:
Db         76 TGTATGGGAGCTCTGTTTTCATGCTATTTCACTTTATTAAATCTTGCAACTGCA 23
```

RESULT 170
BX953370
LOCUS BX953370 462 bp mRNA linear EST 01-MAR-2004
DEFINITION DKFZp781D0546_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone DKFZp781D0546 5', mRNA sequence.
ACCESSION BX953370
VERSION BX953370.1 GI:43432076
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 462)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. No s1 sequence available. This clone (DKFZp781D0546) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES Location/Qualifiers
source 1. .462
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781D0546"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

Query Match 5.4%; Score 404.5; DB 1; Length 462;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 409; Conservative 30; Mismatches 20; Indels 1; Gaps 1;

```
Qy       2786 CCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTT 2845
             || |||||||:|||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          4 CCGGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTT 63

Qy       2846 TGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGG 2905
             |||||||||||| |||||||||:  |||||| ||||||||||:|||||||||||||||||
Db         64 TGCCTTTGAAGATCCTTCAAACCCAATGTCTCAATTCACCTGGACTGTTTTACCCCAAGG 123

Qy       2906 GTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTC 2965
             ||||||||| ||:|||||||||||||||||||||||||||||||||||||:||:|::||
Db        124 GTTCAGGGACAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTC 183

Qy       2966 ATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGA 3025
             | ||||||||||||||||||:|||  ||||||||||||||:||:|||: ||||||
Db        184 ACACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCTGCCCTTTCAGA 243

Qy       3026 AACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||||||||||||||||||||||||||:|||||:|||||||| ||:|||||||||||| :
Db        244 AACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTTGCCACCTGTGGCTACCA 303

Qy       3086 GGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATT 3145
             ||||||||||:|:| |||||:|||||||||||||||| ||||||||||||:||||:|||
Db        304 GGTTTCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAAAATT 363

Qy       3146 ATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCA 3205
             |||| |||||||||:|||||||||||||:::| |||||| | ||||||||||||||||
Db        364 ATCCTAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATCCTCA 423

Qy       3206 TCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTA 3245
             ||:| |||||||||||||||||:|::|||||| ||:||
Db        424 TCCC-AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATA 462
```

RESULT 171
AQ226613
LOCUS AQ226613 479 bp DNA linear GSS 26-SEP-1998
DEFINITION HS_2010_A2_C05_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2010 Col=10 Row=E, genomic survey sequence.
ACCESSION AQ226613
VERSION AQ226613.1 GI:3651842
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 479)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and

```
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2010  row: E  column: 10
            Class: BAC ends
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2010 Col=10 Row=E"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.4%;  Score 402.9;  DB 1;  Length 479;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  405;  Conservative   37;  Mismatches   32;  Indels    1;  Gaps    1;

Qy       3425 AACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCC 3484
             || :| |||||||| ||:||:|||::|||||| | ||||||:||:| ||||:|||||||
Db             2 AATGGNGCAAGACTNTTCTTTATATGTCACAG-GANAAACAGAAATATCTCTAGGAGTCC 60

Qy       3485 TTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATG 3544
             |||||||:|||:|||||:||||||||||||::||||::||||||:|||||||:|||||
Db            61 TTACACAGGTCCGAGGGACGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATG 120

Qy       3545 TAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT 3604
             ||||||||||||||||:|:||||||||||:|||||||| ||||||||||||||::||||||
Db           121 TAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGATGGCAGTAGCAGTCTTAGTAT 180

Qy       3605 CTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||:||||
Db           181 CTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGA 240

Qy       3665 AYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTC 3724
             |::|||||||||||||||||| |||||||||||||||||||||:|||| |||| ||:||
Db           241 ATGGCATACTCACTGCTAAAGAAGACTTGTGGCTGTCAGACAACTGTTTGCTTACATATC 300

Qy       3725 AGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAG 3784
             ||||||||||||||||:||||||||  |:|||| ||||||||||||||||||||||||||
Db           301 AGGCTCTATTACTTGAAAGGCCAGTGCCACGACTGCGCACTTGTGCAACTCTTAACCCAG 360

Qy       3785 YCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTC 3844
             :| |||||||||||||||||| ||| || |:||:|||||||| ||||:||||| |||
Db           361 CCACATTTCTTCCAGACAATGAATAAAGGAAAGAACATAACTGTCCACAAGTAATTGCTC 420

Qy       3845 AAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCT 3899
             | ||||||||   | || | |||  | ||||:||  |:||||||||||: ||||
Db           421 AGACCTATGCTGATNGANGNGACTNTNTAGAGGTCTCCTTGACTGATCCCAACCT 475


RESULT 172
AQ486506
LOCUS       AQ486506                539 bp    DNA     linear   GSS 24-APR-1999
DEFINITION  RPCI-11-243O2.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-243O2,
            genomic survey sequence.
ACCESSION   AQ486506
VERSION     AQ486506.1  GI:4672380
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-243O2.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7593265"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-243O2"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             5.4%;  Score 402.1;  DB 1;  Length 539;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  452;  Conservative   22;  Mismatches   52;  Indels    5;  Gaps    3;

Qy       4154 CTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCW 4213
             ||||||||:|||:||||||||||||:||||||||||||||||||||:|||||||:||||:
Db             1 CTAACTTCTGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCT 60

Qy       4214 GTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAA 4273
             |||||||||||||:|||||||:|||||||||||||:|||||||||| ||||||||:||
Db            61 GTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAAGGAAAGGAA 120

Qy       4274 AGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGAC 4333
             |||||||||:|||:|||:||||||||:||||||||:|||||||||||||||||||||||
Db           121 AGGGAAATAGAAGGGAACTGCCAAGCAGATATTGAAGCCAAAAGAGCTGCAAGGCAGGAC 180

Qy       4334 CCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACC 4393
             | |||||||||||||||| || | ||| |||||| |||||||||| ||| |||||||
Db           181 CTTCCATTAGAAATGCTTA-TAGAGGACCCCTAGTATGGGGTAATCCCCTCCAGGAAACC 239

Qy       4394 AAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTC 4450
             |||||||||||||||| || | ||| |||||| |||||||||| || |    ||||| |||
Db           240 AAGCCCCAGTACTCAGAAGAAAAAATAGAATGAGGAACCTCATGAAGACATAGTTTCCTC 299

Qy       4451 CCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGG 4510
             ||||| ||| || |||||| ||||||| |||||||| |||||||| ||| ||| | |
Db           300 CCCTCAGGATGGCCAGCCACCAAAGAAGGAAAAATACTTATGCCTGCAGCTAACCAGTAG 359

Qy       4511 AAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATG 4570
             |||||||||||||||| | |||||||||||||| ||||| |||||||||||||:|||
Db           360 AAATTACTTAAAACCCTTTACCAAACCTTTCACTTTGGCATTGATAGCACCCATCAGATG 419

Qy       4571 GCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCC 4630
             ||||||| |||||||||| ||||||| | |||||||||||||||| :|||:||||||||
Db           420 GCCAAATTATTATTTACTAGACCAGGTC-TTTCAAAACTATCAAGCGGATAGTCAGGGCC 478
```

```
Qy       4631 TGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCA 4681
              |||||||:|||||||:|:|| |||||||||||  :   | | |     ||||
Db        479 TGTGAAGTGTGCCAAAGAAGTAATCCCCTGCACTGCAGGCCATACATTTCA 529


RESULT 173
BF992535
LOCUS       BF992535                423 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-301000-203-e12 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF992535
VERSION     BF992535.1  GI:12398858
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            301000-203-e12&t3=2000-10-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 408.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.4%;  Score 401.4;  DB 1;  Length 423;
  Best Local Similarity   96.6%;  Pred. No. 0;
  Matches  393;  Conservative   14;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4479 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 4538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 60

Qy       4539 TTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCC 4598
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         61 TTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCC 120

Qy       4599 TTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCC 4658
              ||||||||||||||||||:|||:|||||||||||||:|||||||:|:||||||||||||
Db        121 TTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCC 180

Qy       4659 TGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACT 4718
              |||||:|||||||||||||||||||:||||||:||||||||||||||||||:|:||:|||
Db        181 TGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACT 240

Qy       4719 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 4778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 300

Qy       4779 ARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGT 4838
              |:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        301 AGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGT 360

Qy       4839 AATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 407


RESULT 174
AQ061550/c
LOCUS       AQ061550                557 bp    DNA     linear   GSS 31-JUL-1998
DEFINITION  CIT-HSP-2347F8.TF CIT-HSP Homo sapiens genomic clone 2347F8,
            genomic survey sequence.
ACCESSION   AQ061550
VERSION     AQ061550.1  GI:3363462
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2347F8.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2347F8"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             5.3%;  Score 399.9;  DB 1;  Length 557;
  Best Local Similarity   84.3%;  Pred. No. 0;
  Matches  424;  Conservative   43;  Mismatches   25;  Indels   11;  Gaps    3;

Qy       2921 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCT 2980
              |||||||||||||||||||||| |||||||||||:||:|::||||||| |||||||||
Db        557 CCATCTATTTGGCCAGGCATTAGCNCAAGACTTGAGCCAGTTCTCATACCAGGACACTCT 498

Qy       2981 TGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCA 3040
              ||||||||:|||  |||||||||||||||:||:|||::|||||||||||||||||||||
Db        497 TGTCCTTCAGTACATGGATGATTTACTTTTAGCCGCCTGTTCAGAAACCTTGTGCCATCA 438
```

```
Qy       3041 AGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARA 3100
Db        437 AGCCACCCGAGCGCTCTTAAATTTCCTTGCCACCTGT--------GGTTTCCAAACCAAA 386

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
Db        385 GGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGTACCAG 326

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
Db        325 GGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCGAAACCCTAAC 266

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
Db        265 GCGACTAAGAGGGTTCTTTGGCATAACAGGCTTCTGCTGAATATGGATT-CCCAGGTATG 207

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
Db        206 GCAAAATAGCCAAGCCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACTCATT 147

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
Db        146 TAGTAAGATGGACACCTGAAGCAGAATCGGCTTTCTAGGCCCTAAAGAAGGTCCT--AAC 89

Qy       3401 CCAAGYCCCAGTGTTAAGYTTGC 3423
Db          88 CCAAGCCACAATGTTAAGCTTGC 66
```

```
RESULT 175
AW967582
LOCUS       AW967582                529 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST379657 MAGE resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW967582
VERSION     AW967582.1  GI:8157419
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 243
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGJ"
                     /note="Vector: pBluescriptSKm"

  Query Match             5.3%;  Score 399.2;  DB 1;  Length 529;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  410;  Conservative   29;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       3968 AGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCA 4027
Db          32 AGGAATAAAGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCA 91

Qy       4028 GAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATAT 4087
Db          92 GAACTAATAACCCTCACTCGGGCACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATAT 151

Qy       4088 ACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
Db         152 ACAGACTCTAAGTCTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGG 211

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
Db         212 GAATTCCTAATTTCCGAGGGAACACCTATAGAACATCAGGAAGCCATTAGGAGATTACTA 271

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
Db         272 TTGGCTATACAGAAACCTAAAGAGGTGGCAGACTTACACTGTTGGGGTCACCAGAAAGGA 331

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
Db         332 AAGGAAAGGGAAATAGAAAGGAACCGCCAAGCGGATCTTAAAGTCAAAAGAGCCGCAAGG 391

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
Db         392 CAGGACCCTCCATTAGAAATGCTTATGGAAGGGCCCCTAGTACAGGGTAATCCCCTCCGG 451

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGGCAGTTTT 4447
Db         452 GAAACCAAGCCCCAGTTTTCAGCAGGAAAAATAGAATTAGGAACCTTACAAGGACATACC 511

Qy       4448 CTCCCCTC 4455
Db         512 TTCCTCCC 519
```

```
RESULT 176
AI128496/c
LOCUS       AI128496                415 bp    mRNA    linear   EST 27-OCT-1998
DEFINITION  qc61e08.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1714118 3' similar to contains PTR7.t1 PTR7 repetitive
            element ;, mRNA sequence.
ACCESSION   AI128496
VERSION     AI128496.1  GI:3597010
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 415)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1184   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .415
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1714118"
                     /dev_stage="two placentae: one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATTTTTTTTTTTTTTTTT 3'],
```

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

Query Match 5.3%; Score 397.3; DB 1; Length 415;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 403; Conservative 9; Mismatches 3; Indels 1; Gaps 1;

```
Qy       7166 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAG 7225
             ||||||||||||||||||||||||||||||||||||||:||||||||| |||| || |||
Db        415 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAA-CTCCNCANCAG 357

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 297

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             ||||:|||||||||:|||||||||:||||||||||||||||||||||||||||||||||||
Db        296 CTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 237

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 7405
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 177

Qy       7406 GACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAAT 7465
             |||||||||||||||||||||||||||||:||||||:||||||||||||||||||||||||
Db        176 AACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAAT 117

Qy       7466 GATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCC 7525
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        116 GATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCC 57

Qy       7526 CTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          56 CTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1
```

RESULT 177
AQ483208
LOCUS AQ483208 534 bp DNA linear GSS 24-APR-1999
DEFINITION RPCI-11-213C19.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-213C19, genomic survey sequence.
ACCESSION AQ483208
VERSION AQ483208.1 GI:4670612
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 534)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-213C19.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .534
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7581474"
/db_xref="taxon:9606"
/clone="RPCI-11-213C19"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

Query Match 5.3%; Score 395.2; DB 1; Length 534;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 443; Conservative 20; Mismatches 64; Indels 4; Gaps 2;

```
Qy       4154 CTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCW 4213
             ||||| ||:|||:||||||||||:||||||||||||||||||:||||||:|||
Db          1 CTAACATCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTG 60

Qy       4214 GTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAA 4273
             ||||| |||||||:||| ||||:||||||||||||:|||||||||| |||||| ||:||
Db         61 GTACAAAAACCTAAAGAAGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAGAGAAA 120

Qy       4274 AGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGAC 4333
             |||||||| |||: ||: |||||| : |||| ||||:|||||||||||| |||||||
Db         121 AGGGAAATAAAAGAAAACCGCCAAGTGGTTATTAAAGCAAAAAGAGCTGCAA-GCAGGAC 179

Qy       4334 CCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACC 4393
             |||||||||||||||||||  ||   ||| |||||| |||||||||| ||| |||| ||
Db         180 CCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCAGGAAGCC 239

Qy       4394 AAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTC 4450
             |||||||||||||||||  ||||| |||||| |||||| ||||||    ||||| |||
Db         240 AAGCCCCAGTACTCAGCATAAGAAATAGAATGGAGAACCTTACGAGGACATAGTTTCCTC 299

Qy       4451 CCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGG 4510
             ||||| ||| || |||||||  ||||||| ||| |||||||||||||| ||| |||||||
Db         300 CCCTCAGGATGGCTAGCCACCAAAGAAGGAAAAGTACTTTTGCCTGCAGCTAACCAATGG 359

Qy       4511 AAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATG 4570
             |||||||||||||||||| |||||||||||||| |||| ||||||||||||||:|||
Db         360 AAATTACTTAAAACCCTTCACCAAACCTTTCACTTATGCATTGATAGCACCCATCAGATG 419

Qy       4571 GCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCC 4630
             | ||||  |||||||||||||||||||||||||||||||||||||||:|||: ||||| |
Db         420 GTCAAAGTGTTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGACAGGGAC 479

Qy       4631 TGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCA 4681
             | ||||:| |||||:|:|| |||||||| |  :|  | | |     ||||
Db         480 TATGAAGTATGCCAAAGAACTAATCCCCTACACTACAGGCCATACATTTCA 530
```

RESULT 178
BI963185/c
LOCUS BI963185 543 bp mRNA linear EST 12-MAR-2002
DEFINITION ie57c10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens cDNA clone IMAGE:5670930 3', mRNA sequence.
ACCESSION BI963185
VERSION BI963185.1 GI:16337590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,

```
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu) This sequence now available from the IMAGE
            consortium, for clone orders contact: info@image.llnl.gov
            High quality sequence stop: 451.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5670930"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
                     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
                     Site_2: Sal 1; Starting library constructed using
                     SuperScript Plasmid Library kit (Life Technologies). cDNA
                     made by oligo-dT priming. Size-selected by column
                     fractionation; average insert size 1.08 kb. Library was
                     amplified once on solid support and plasmid DNA from
                     library was prepared. The library DNA was normalized by
                     method #4 from Bonaldo, Lennon, and Soares 1996 Genome
                     Research 6:791-806; 0.5 microgram single-stranded library
                     plasmid DNA was mixed with 5 micrograms PCR product
                     representing library inserts and hybridized to an Ecot of
                     20. Single-stranded (unhybridized) plasmids were isolated
                     by hydroxyapatite chromatography and used to make this
                     library."

  Query Match             5.3%;  Score 395.1;  DB 1;  Length 543;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  457;  Conservative    8;  Mismatches   53;  Indels   11;  Gaps    3;

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  ||||||||||||  |||||||||||||  ||  ||||  ||||||||||||
Db        543 AAATGGAGCCCCAGATGCAGTCCATGACTAAGATCTACTGCGGACCTCTGGACCGGCCTG 484

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||  |  ||   |  |||||||||||||||||||||||||  ||||||||||| |||
Db        483 CTAGCCCA-GCTCCCACATTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTG 425

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||  |||||||||||||||||||||||||||||||||||  |||:|| |||||||||
Db        424 CATGACACCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTCGTTGGCCAACCT 365

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||| |||||| |||||||||||||||||||||||||||||||||||||||||
Db        364 CCCCAACAGCACTTGGGTTTT-CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 306

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||  |:|||||||:|||||||:|||||||||||| ||| |||||||||
Db        305 TTTCCTAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTA 246

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAAT---------CAGAGAGCTCACTA 7387
             | |||  |||||||||||||:|||||| ||||| |         ||||||||  |||
Db        245 ACCACTAGGCTTGCAACTTAGCTCACACCCGACCANTAGGGTAGTAAAGAGAGCTTGCTA 186

Qy       7388 AAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAG 7447
             |||||||||||||||||| |||||||||||||||||||||| ||||:||| ||:||| ||
Db        185 AAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAG 126

Qy       7448 CACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAAC 7507
             ||||   || |||||||||||| |||||||||| || |  ||| |||| |||| ||| ||
Db        125 CACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTCAAGCCAGCAATGGCTAC 66

Qy       7508 CCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATT 7556
             || |||||||||||||||| || ||||||||||||||||||  |||||
Db         65 CCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCACTCTATT 17


RESULT 179
DB311329/c
LOCUS       DB311329                 527 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB311329 BRTHA2 Homo sapiens cDNA clone BRTHA2012272 3', mRNA
            sequence.
ACCESSION   DB311329
VERSION     DB311329.1  GI:83149411
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRTHA2012272"
                     /tissue_type="thalamus"
                     /clone_lib="BRTHA2"
                     /note="Vector: pME18SFL3"

  Query Match             5.3%;  Score 394.1;  DB 1;  Length 527;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  449;  Conservative    9;  Mismatches   33;  Indels   21;  Gaps    3;

Qy       7080 ACCCCTGGACCGGCCTGYTAGCCCAC--GATCTGATGTTAATGACATCAAAGGCACCCCT 7137
             ||| ||||||| |||||:||||||  | || |||||||||||||| |||||||| |
Db        502 ACCTCTGGACCAGCCTGCTAGCCCACATGCTCCGATGTTAATGACATCGAAGGCACCCGT 443

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             |  |||||||||||| |||||||||| ||||||| |||||||||||||||||||||||||
Db        442 CTCGAGGAAATCTCAACTGCACAACCCCTACTACACCCCAATTCAGCAGGAAGCAGTTAG 383

Qy       7198 AGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTG 7257
```

```
                ||||||:||| |||||||||||| |||||||| ||||||||||||||||  ||||||||
Db        382 AGCGGTCGTCAGCCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGCGGGGACTG 323

Qy       7258 AGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAG 7317
                |||||||||||||||||||||||||||||||:|||||||||:|||||||||:|||||||
Db        322 AGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAG 263

Qy       7318 GTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC--- 7374
                |||||| | || |||||||||||||||||||||||||||:|||||||  |||||||
Db        262 GTGACCGCGTCTACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGA 203

Qy       7375 ------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCA 7428
                     |||||||||||||||||||||||||||||| |||||||||| ||||||||||
Db        202 TAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAGAGAAATAGCCA 143

Qy       7429 ATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYT 7488
                |||||:||| ||:||||||||||||||||||||||||||||| ||||||||||||| |
Db        142 ATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCCGGATATAAACCCAAGCAT 83

Qy       7489 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTC 7548
                |||||||  |||| ||| |||| ||||| |||||||||||||||||||||||||| |
Db         82 TCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTGTATGGGAGCTCTAT---- 27

Qy       7549 ATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
                |||||||||||||||||||||||:|||
Db         26 ------TTTCACTCTATTAAATCTTGCAGCTG 1


RESULT 180
AQ141344
LOCUS       AQ141344                516 bp    DNA     linear   GSS 24-SEP-1998
DEFINITION  HS_3150_A1_G03_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3150 Col=5 Row=M, genomic survey
            sequence.
ACCESSION   AQ141344
VERSION     AQ141344.1  GI:3531997
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3150  row: M  column: 5
            Class: BAC ends
            High quality sequence stop: 516.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3150 Col=5 Row=M"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.3%;  Score 393.3;  DB 1;  Length 516;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  421;  Conservative   17;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4179 MACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAG 4238
                :|||||||||||||||||||||:|||| ||:||||:|||||||||| ||:||||||:|
Db         31 CACATCAGGAAGCCATTAGGAGATTACTATTGGCTGTACAGAAACATAAAGAGGTGGCAA 90

Qy       4239 TCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAG 4298
                |||| | ||||:|| |||| || |||||||||:||||||||||:|| :|||:  |||||
Db         91 TCTTTCTCTGCCGGTGTCACCAGAAAGGAAAGGAAAGGGAAATAGAAAGGAACCACCAAG 150

Qy       4299 CAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAAC 4358
                | :|||||||||:|||||||| ||||||||||||||||||||||||||||||  ||
Db        151 CGGATATTGAAGCCAAAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAAG 210

Qy       4359 TTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAA 4418
                  ||| |||||| |||||||||| ||| |||||||||||||||||||||| || |||||
Db        211 GACCCCTAGTATGGGGTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGAAGAAGAAA 270

Qy       4419 CAGAATGGGGAACCTCACGAG---GCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAG 4475
                |||||| |||||||||| ||     ||||| |||||||| ||| || ||||||||||||
Db        271 TAGAATGAGGAACCTCACAAGCACATAGTTTCCTCCCCTCAGGATGGCTAGCCACTGAAG 330

Qy       4476 AAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAA 4535
                |||| |||||||||||||||||| ||| |||||||||||||||||||||||||| ||||
Db        331 AAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAAA 390

Qy       4536 CCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAG 4595
                | |||| |||||||||| ||||||||||||||:|||||||||| ||||||||||||||||
Db        391 CATTTCCCTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATTTACTGGACCAG 450

Qy       4596 GCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATC 4655
                |||||||||||||||||||||:|||:|||||||||||| ||:||||||| | || |||||
Db        451 GCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTAAAGTGTGCCATACAAGTAATC 510

Qy       4656 CCCTGC 4661
                ||||||
Db        511 CCCTGC 516


RESULT 181
N53177/c
LOCUS       N53177                  422 bp    mRNA    linear   EST 28-JAN-1997
DEFINITION  yv56h11.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:246789 3' similar to contains PTR7.t2 PTR7 repetitive element
            ;, mRNA sequence.
ACCESSION   N53177
VERSION     N53177.1  GI:1194343
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
           Insert Length: 938   Std Error: 0.00
           Seq primer: m13 -40 forward
           High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3796035"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:246789"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             5.2%;  Score 392.9;  DB 1;  Length 422;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  409;  Conservative   10;  Mismatches    2;  Indels    3;  Gaps    2;

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCC 7218
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        422 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCC 363

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 303

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             ||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||||||||
Db        302 TCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAA 243

Qy       7339 CACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATT 7398
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        242 CACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATT 183

Qy       7399 AGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAG 7458
             ||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||
Db        182 AGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAG 123

Qy       7459 GGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGT 7518
             ||||||||||||||||||||||||||||||:|||||  ||||||||||||||||  || |||
Db        122 GGACAATGATCGGGATATAAACCCAAGTCTTCGA-NCGGCAACGGCAACCCC--TTTGGT 66

Qy       7519 CCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARC 7578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db          65 CCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAAC 6

Qy       7579 TGCR 7582
             |||:
Db           5 TGCA 2


RESULT 182
AG193108
LOCUS       AG193108               501 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-070A05.T7, genomic survey
            sequence.
ACCESSION   AG193108
VERSION     AG193108.1  GI:45225284
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 501)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-070A05.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"

  Query Match             5.2%;  Score 392.9;  DB 1;  Length 501;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  420;  Conservative   18;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4179 MACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAG 4238
              :|||||||||||||||||||:| | |||:||||:|||||||||||||:||||||||:||
Db          1 AACATCAGGAAGCCATTAGGAGAGTGTTATTGGCTGTACAGAAACCTAAAGAGGTGGCAG 60

Qy       4239 TCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAG 4298
              ||||| |||||:|||||||||| |||||||| :||| |||||||:||| |||: ||||||
Db         61 TCTTATACTGCTGGGGTCATCAGAAAGGAAATGAAAAGGAAATAGAAGTGAACCGCCAAG 120

Qy       4299 CAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAAC 4358
              | :||| ||||||:|||||| ||||||||| |||||||| |||||||||||||||  ||
Db         121 CGGATACTGAAGCCAAAAGACCTGCAAGGCGGGACCCTCTATTAGAAATGCTTATAGAAA 180

Qy       4359 TTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAA 4418
                 ||| |||||| |||||||||| ||| |||||||||||||||||||||||||| ||| |
Db         181 GACCCCTAGTATGGGGTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAGA 240

Qy       4419 CAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAG 4475
              |||||||||||||||| ||||    ||||| |||||||| ||| || |||||||  |||
Db         241 TAGAATGGGGAACCTCAGGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAAG 300

Qy       4476 AAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAA 4535
              |||| |||||||||||||| ||| ||| |||||||||||||||||||||||| ||| ||||
Db         301 AAGGAAAAATACTTTTGCCGGCAGCTAACCAATGGAAATTACTTAAAACCCATCACCAAA 360

Qy       4536 CCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAG 4595
              |||||||||||||||| ||||||||||||:|| |||||||||||||||||||||||||
Db         361 CCTTTCACTTAGGCATTGATAGCACCCATCAGATAGCCAAATCATTATTTACTGGACCAA 420

Qy       4596 GCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATC 4655
              |||||||||||||||||||||:|||:|||||||||||| ||:|||||||:|:||| ||||
Db         421 GCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTAAAGTGTGCCAAAAAAAAAATC 480

Qy       4656 CCCTGC 4661
              | ||||
Db         481 CACTGC 486
```

```
RESULT 183
AQ231407/c
LOCUS       AQ231407                549 bp    DNA     linear   GSS 26-SEP-1998
DEFINITION  HS_2027_B2_H08_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2027 Col=16 Row=P, genomic survey
            sequence.
ACCESSION   AQ231407
VERSION     AQ231407.1  GI:3656636
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2027  row: P  column: 16
            Class: BAC ends
            High quality sequence stop: 549.
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2027 Col=16 Row=P"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.2%;  Score 392;  DB 1;  Length 549;
  Best Local Similarity   82.9%;  Pred. No. 0;
  Matches  408;  Conservative   42;  Mismatches   38;  Indels    4;  Gaps    2;

Qy       2929 TTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTC 2988
             |||||||||||||||||||||||| ||||:||:|::|| ||||||||||||||||||||
Db        528 TTGGCCAGGCATTAGCCCAAGACNTGAGCCAATTCTCCTACCTGGACACTCTTGTCCTTC 469

Qy       2989 RGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCC 3048
             :||| |||||||| |||||| :||: ||::||||||| | |||||||||||||||||| |
Db        468 GGTACATGGATGATNTACTTTAAGCCACCCATTCAGAACCTTTGTGCCATCAAGCCACGC 409

Qy       3049 AAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARC 3108
             ||||:|| ||:|||||||| ||:|||||||||||||:||||||||||:|:|:|||||:|
Db        408 AAGCACTATTAAATTTCCTTGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGC 349

Qy       3109 TCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
             |||||||||||||||||||||||||||||||||||:||| :|||||||||||||||||:  ||||||
Db        348 TCTGCTCACAGCAGGTTAAATACTTAGGGCTANAATTATCCAAAGGCACCAGAACCCTCA 289

Qy       3169 GTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAA 3228
             ||||||||:::|||||||||||||||||| ||||||||||:|||||||| ||||||||||
Db        288 GTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCATAAAGCAACTAA 229

Qy       3229 GRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATA 3288
               :| :|||||||||:|||:|||:||||||| ||:||||||| | |||||:  ||: ||||
Db        228 CAGCGTTCCTTGGCATAACAGGTTTCTGCCAAATATGGATT-CNCAGGTACAGCANAATA 170

Qy       3289 GCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGA 3348
             ||||| :|||||:|||||: |||||||||||||| ||||||||||||||||:|||||
Db        169 GCCAGACCATTAAATACACGAATTAAGGAAACTCANAAAGCCAATACCCATTTAGTAAGA 110

Qy       3349 TGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
             ||||:|:|||||||:::|||||||||||||||||||  ||  | |  || | ||   |  ||
Db        109 TGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCGCTAACAC---ATGCC 53

Qy       3409 CAGTGTTAAGYT 3420
             ||||||||||:|
Db         52 CAGTGTTAAGCT 41


RESULT 184
R76086/c
LOCUS       R76086                  458 bp    mRNA    linear   EST 06-JUN-1995
DEFINITION  yi71b03.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:144653 3' similar to contains PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   R76086
VERSION     R76086.1  GI:850768
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 597
            High quality sequence stops: 337
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 597   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 337.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:562655"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:144653"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             5.2%;  Score 391.9;  DB 1;  Length 458;
  Best Local Similarity   92.5%;  Pred. No. 0;
```

```
  Matches  406;  Conservative    8;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||| ||||| ||||          | ||||||||||||||||||||||| ||
Db        438 AAGGCACCCNTCCTGGGGAANCTCAGCTGCAAAAACCTCTACTACGCCCCAATTCAGNAG 379

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        378 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 319

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        318 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 259

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        258 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCT 199

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GACCAATCAGAGAGCTCACTNAAATGCTAATTAGGCAAAGANNGGAGGTNAAGAAATAGC 139

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:|||||||:|||||| |||||||||||||||||| |||||||||||||||||||
Db        138 CAATCATCTATTGCCTGAGAGGACAGCAGGAGGGACAATGANCGGGATATAAACCCAAGT 79

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :|||||||||||||||| | ||||||||||||||||||| ||||||||||||||||||||
Db          78 CTTCGAGCCGGCAACGGCNA-CCCCTTTGGGTCCCCTCCCCTTGTATGGGAGCTCTGTTT 20

Qy       7547 TCATGCTATTTCACTCTAT 7565
             |||||||||||| |||||||
Db          19 TCATGCTATTTNACTCTAT 1


RESULT 185
DB297732/c
LOCUS       DB297732                542 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB297732 BRACE2 Homo sapiens cDNA clone BRACE2027568 3', mRNA
            sequence.
ACCESSION   DB297732
VERSION     DB297732.1  GI:83106148
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.

FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRACE2027568"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE2"
                     /note="Vector: pME18SFL3"

  Query Match             5.2%;  Score 391.9;  DB 1;  Length 542;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  448;  Conservative    8;  Mismatches   35;  Indels   21;  Gaps    3;

Qy       7080 ACCCCTGGACCGGCCTGYTAGCCCAC--GATCTGATGTTAATGACATCAAAGGCACCCCT 7137
             ||| ||||||| |||||:||||||||  | || ||||||||||||||| |||||||| |
Db        502 ACCTCTGGACCAGCCTGCTAGCCCACATGCTCCGATGTTAATGACATCGAAGGCACCCGT 443

Qy       7138 CCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             |  |||||||||||| |||||||||| ||||||| |||||||||||||||||||||||||
Db        442 CTCGAGGAAATCTCAACTGCACAACCCCTACTACACCCCAATTCAGCAGGAAGCAGTTAG 383

Qy       7198 AGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTG 7257
             |||||| ||| |||||||||||||| |||||||| ||||||||||||||  ||||||||
Db        382 AGCGGTNGTCAGCCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGCGGGGACTG 323

Qy       7258 AGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAG 7317
             ||||||||||||||||||||||||||||||:|||||||||:|||||||||:|||||||
Db        322 AGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAG 263

Qy       7318 GTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC--- 7374
             |||||| | || |||||||||||||||||||||||||||||:|||||||  |||||||
Db        262 GTGACCGCGTCTACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGA 203

Qy       7375 ------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCA 7428
                   |||||||||||||||||||||||||||||| |||||||||| ||||||||||
Db        202 TAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAGAGAAATAGCCA 143

Qy       7429 ATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYT 7488
             |||||:||| ||:|||||||||||||||||||||||||| ||||||||||||||  |
Db        142 ATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCCGGATATAAACCCAAGCAT 83

Qy       7489 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTC 7548
             ||||||| |||| ||| |||| ||||| |||||||||| ||||||||||||||| |
Db         82 TCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCCTTGTATGGGAGCTCTAT---- 27

Qy       7549 ATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
                  |||||||||||||||||||:|||
Db         26 ------TTTCACTCTATTAAATCTTGCAACTG 1


RESULT 186
DB375050/c
LOCUS       DB375050                542 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB375050 PLACE1 Homo sapiens cDNA clone PLACE1004321 3', mRNA
            sequence.
ACCESSION   DB375050
VERSION     DB375050.1  GI:83095124
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
```

```
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004321"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             5.1%;  Score 385.2;  DB 1;  Length 542;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  463;  Conservative    7;  Mismatches   58;  Indels   10;  Gaps    5;

Qy       7044 GCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCC 7103
             | |||||| |||||||||| ||      |  || ||| |||||||||| |||:||||||
Db        528 GGCCAAGAGGCAGTCCAAGATTA--GATTTACGNAGAACCCTGGACCGGGCTGTTAGCCC 471

Qy       7104 ACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACC 7163
             | ||| ||||||||    |     || |   || ||| | |||| | |  |    | |||
Db        470 AAGATTTGATGTTANGAAT---NAAAGGNACCTTCCGGNGGAATTTTAGGTG--AAAACT 416

Qy       7164 TCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAAC 7223
             | ||  | ||||||||| ||||||||||  |||||| ||| || |||||||||| |||||
Db        415 TTTATAAGGCCCCAATTTAGCAGGAAGCNATTAGAG-GGTTGTTGGCCAACCTTCCCAAA 357

Qy       7224 AGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTA 7283
                | |||||||||||||||||||   |||||||||||| ||| ||||||||||
Db        356 NAGNAATTAGTTTTCCTGTTGAGATGGGG--ATGAGAGACAGGANTAGNTGGATTTCCTA 299

Qy       7284 GGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGG 7343
             |||||| |||||||||:|||||||||:||||||||||||||||||||||||||||||||||
Db        298 GGCTGANTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGG 239

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 7403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GGCTTGCAACTTAGNTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 179

Qy       7404 AAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACA 7463
             |||||||||||||||||||||||||||||:||||||:||||||||||||||||||||||||
Db        178 AAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACAGCAGGAGGGACA 119

Qy       7464 ATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCT 7523
             |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        118 ATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCT 59

Qy       7524 CCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||||||||||||||||||||||| |||||||||||||:||||
Db         58 CCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTNTATTAAATCTTGCAACTGC 1


RESULT 187
R68740
LOCUS       R68740                 459 bp    mRNA    linear   EST 26-JUN-1995
DEFINITION  yi14g06.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:139258 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
            POLYPROTEIN - ;, mRNA sequence.
ACCESSION   R68740
VERSION     R68740.1  GI:842257
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 459)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1644
            High quality sequence stops: 316
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1644   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 316.
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:545824"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:139258"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             5.1%;  Score 384.4;  DB 1;  Length 459;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  433;  Conservative    3;  Mismatches   19;  Indels    4;  Gaps    4;

Qy       6193 CAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTT 6252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTT 60

Qy       6253 TCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACA 6312
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACA 120

Qy       6313 TACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTA 6372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTA 180

Qy       6373 CCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCA 6432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCA 240

Qy       6433 GAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGAT 6492
             |||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db        241 GAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGAT 300
```

```
Qy       6493 TTATACAGTTATGTCATATCTAA-GCCCCGCAACAAAAGAGTACCCA-TTCTTCCTTTTG 6550
             ||||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||
Db        301 TTATACAGTTATGTCATATCTAAGGCCCCGCAACAAAAGAGTACCCATTTCTTCCTTTTG 360

Qy       6551 TTATA-GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCT 6609
             |||||  |||||||| |||  | ||||||  | ||||  |||| |  |||||||||||||||  |
Db        361 TTATAGGGAGCAGGGGTGTTGGGTGCATTGGGTANTGGCNTGGGCGGTATCACAACCTTT 420

Qy       6610 ACTCAGTTC-TACTACAAACTATCTCAAGAACTAAATGG 6647
             | ||||||  || ||||||  | |||||  | |  |||  ||
Db        421 ATTCAGTTNTTATTACAAANTTTCTCAGGGANTAATGGG 459


RESULT 188
BG004254/c
LOCUS       BG004254                411 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  RC6-GN0070-301100-014-C06 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG004254
VERSION     BG004254.1  GI:12445242
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            301100-014-C06&t3=2000-11-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 21
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.1%;  Score 384;  DB 1;  Length 411;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches  381;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        411 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAGCTATCTCAAGAACTAAA 352

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 292

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGG 6764
             ||||||||||||||||:||||||||||||||||||:||||||||||||:||||||||
Db        291 AGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGG 232

Qy       6765 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 6824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 172

Qy       6825 TGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACAC 6884
             ||||||||||:|||||||||:||||||||||||||||:||||||:|||||||||||||||
Db        171 TGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACAC 112

Qy       6885 TGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCT 6944
             ||||||||||||||||||||||:|||||||||||||||||||||||| ||||||| |||
Db        111 TGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTATTAGGACTTCT 52

Qy       6945 AGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||
Db          51 AGCAGCTATAATATTGCTACTCCTCTTTGGAC 20


RESULT 189
AW511366/c
LOCUS       AW511366                485 bp    mRNA    linear   EST 03-MAR-2000
DEFINITION  hd45h03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2912501 3' similar to contains L1.t2 L1 repetitive element ;,
            mRNA sequence.
ACCESSION   AW511366
VERSION     AW511366.1  GI:7149444
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2912501"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-PacI; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (fetal lung NbHL19W, testis NHT, and
                     B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made
                     in vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "

  Query Match             5.1%;  Score 382.7;  DB 1;  Length 485;
  Best Local Similarity   89.0%;  Pred. No. 0;
```

```
  Matches  414;  Conservative    7;  Mismatches   35;  Indels    9;  Gaps    1;

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160
              | || | || |||||||||||||| ||||||||| ||  ||||||||||| |||||||
Db        480 CACATGCTCCGATGTTAATGACATCGAAGGCACCCGTCTCGAGGAAATCTCAACTGCACA 421

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220
              ||| ||||||| ||||||||||||||||||||||||||||||:||| ||||||||||||
Db        420 ACCCCTACTACACCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCAGCCAACCTCCCC 361

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280
              | |||||||| ||||||||||||||||  ||||||||||||||||||||||||||||||
Db        360 AGCAGCACTTGGGTTTTCCTGTTGAGAGCGGGGACTGAGAGACAGGACTAGCTGGATTTC 301

Qy       7281 CTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACA 7340
              |||||||||:|||||||||:|||||||||:|||||||||||||| | || ||||||||||
Db        300 CTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCTACCTTTAAACA 241

Qy       7341 CGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAAT 7391
              |||||||||||||||||:|||||||  |||||||         ||||||||||||||||
Db        240 CGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGATAGTAAAGAGAGCTCACTAAAAT 181

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
              |||||||||||||| |||||||||| |||||||||||||||||:||| ||:||||||||||
Db        180 GCTAATTAGGCAAAAACAGGAGGTAGAGAAATAGCCAATCATCTATCGCCTGAGAGCACA 121

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
              |||||||||||||||| |||||||||||||  |||||||| |||| ||| |||| |
Db        120 GCAGGAGGGACAATGATCCGGATATAAACCCAAGCATTCGAGCCAGCAATGGCTACCCTC 61

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATT 7556
              |||| |||||||||||||||||||||||||||||||| ||||||  |||||
Db         60 TTTGTGTCCCCTCCCTTTGTATGGGAGCTCTATTTTCACTCTATT 16


RESULT 190
N77302
LOCUS       N77302                 433 bp    mRNA    linear   EST 28-JAN-1997
DEFINITION  yv43e03.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:245500 5' similar to PIR:B44282 B44282 retrovirus-related env
            polyprotein - human ;contains PTR5.t1 PTR5 repetitive element ;,
            mRNA sequence.
ACCESSION   N77302
VERSION     N77302.1  GI:1239880
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 907   Std Error: 0.00
            Seq primer: reverse ET
            High quality sequence stop: 305.
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3794746"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:245500"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             5.1%;  Score 379.8;  DB 1;  Length 433;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  417;  Conservative    9;  Mismatches    4;  Indels    4;  Gaps    4;

Qy       6778 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAA 6837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db          1 TTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAA 60

Qy       6838 GAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGC 6897
              |||||||:|||||||||||||||||:||||||:||||||||||||||||||||||||||||
Db          61 GAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGC 120

Qy       6898 CTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 6957
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATA 180

Qy       6958 TTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGA 7017
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         181 TTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGA 240

Qy       7018 ATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 300

Qy       7078 AGACCCCTGGACCGGCCTGYTAGCCCACGATCT-GATGTTAATGACATCAAAGGCACCCC 7136
              ||||||||||||||||||:|||||||||||||| |||||||||||||| ||||| |||
Db         301 AGACCCCTGGACCGGCCTGCTAGCCCACGATCTNGATGTTAATGACATCANAGGCA-CCC 359

Qy       7137 TCCTGAGGAAATCTCAGCT-GCACAACCTCTACTACGCCCCAATTCA-GCAGGAAGCAGT 7194
              |||||||||||||||||| |||||||||||||||||||||||||| ||||| ||||||
Db         360 TCCTGAGGAAATCTCAGCTGGCACAACCTCTACTACGCCCCAATTCAGGCAGGGAGCAGT 419

Qy       7195 TAGAGCGGTSGTCG 7208
              |||||| ||:|| |
Db         420 TAGAGCNGTCGTNG 433


RESULT 191
DQ033443
LOCUS       DQ033443               441 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens HC3142 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ033443
VERSION     DQ033443.1  GI:66884652
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
```

```
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 441)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .441
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     gene            <1. .>441
                     /locus_tag="HC3142"

  Query Match             5.1%;  Score 378.9;  DB 1;  Length 441;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  390;  Conservative   22;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
              || :|||||||||||||||||| |||||||| |:|||||::|||||||||||||||||||||
Db             6 GGAAAGAGATCTTACTGTGTGGCCATCTCATAATGTGAACAGCATACTCACTGCTAAAGG 65

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
              |||||||||||||||||||||: ||| |||||||:|||||||||||||||||||:|||||
Db            66 AGACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCC 125

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||||||||||||||||||||||||||||:| |||||||||||||||||||
Db           126 AGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 185

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||: |||||||||||||:||||| |||||||||| ||  ||||||||||
Db           186 AGAAAAGATAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTGCTCGAGGGGA 245

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||| |||||||:|||| ||||||||| |||||||||||| ||||
Db           246 CCTTCTAGAGGTTCCCTTGGCTGATCCCGACC-TCAACTTGTGTACTGATGGAAGCTCCT 304

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||||| ||| |:|| |||||||||||||||||| |||||||||:|||
Db           305 TTGTAGAAAAAGGACTTCAAAAGGCGGAGTATGCAGTGGTCAGTGACAATGGAATACTTG 364

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||||||||||||| |||:|| |:|||||||||||||:|||:|
Db           365 AAAGTAATCCCCTCACTCCAGGAACTAGCGCTCAGGTGGCAGAACTAATAGCCCTCATTC 424

Qy       4047 GGGCACTAGAATTAGGA 4063
                | ||||||| ||||||
Db           425 AGACACTAGAGTTAGGA 441


RESULT 192
BI052569/c
LOCUS       BI052569                382 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-G04 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052569
VERSION     BI052569.1  GI:14460099
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-G04&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 381.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.0%;  Score 375.6;  DB 1;  Length 382;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  375;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5991 CSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 6050
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            380 CCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 321

Qy       6051 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGA 6110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db            320 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATAA 261

Qy       6111 GGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATA 6170
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||||||
Db            260 GGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCTTGAACTTCAGGCCATA 201

Qy       6171 TGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTC 6230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            200 TGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTC 141

Qy       6231 CGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTG 6290
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db            140 CGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTG 81

Qy       6291 TGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCC 6350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             80 TGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCC 21

Qy       6351 TCCCACACAAATAGTCTGCC 6370
              ||||||||||||||||||||
Db             20 TCCCACACAAATAGTCTGCC 1
```

```
RESULT 193
B38783
LOCUS       B38783                   466 bp    DNA     linear   GSS 18-OCT-1997
DEFINITION  HS-1048-A2-E03-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 831 Col=6 Row=I, genomic survey
            sequence.
ACCESSION   B38783
VERSION     B38783.1  GI:2543035
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 831  row: I  column: 6
            Class: BAC ends
            High quality sequence stop: 466.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 831 Col=6 Row=I"
                     /sex="M"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.0%;  Score 375.5;  DB 1;  Length 466;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  374;  Conservative   34;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       3416 AAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTC 3475
             |||:||||||||:|| ||||||||||:||:||| :|||||| |||||||||:||:|||||
Db         43 AAGCTTGCCAACGGGACAAGACTTTTCTTTATAAGTCACAG-AAAAAACAGGAATAGCTC 101

Qy       3476 TRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGG 3535
             |:|||||||||||||||:|||:||||||:||||||||||||::||||::||||||:|||||
Db        102 TAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCCATGGCATACCTGAGTAAGG 161

Qy       3536 AAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
             |||:|||||||||||||| ||||:|: |||||||||:||||| ||| | |||||||||
Db        162 AAACTGATGTAGTGGCAAAGAGTTGGCCACATTGTTTACGGGTATTGGCGACAGTAGCAG 221

Qy       3596 TYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTC 3655
             |::|||||||||||  |||||||| |||||||:|||||||||||||||||||||||||||
Db        222 TCTTAGTATCTGAAAAAGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTC 281

Qy       3656 ATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTAC 3715
             ||||:||||||::||||||||||||||||||||||||||||||||||||||:||||||
Db        282 ATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTAC 341

Qy       3716 TTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTC 3775
             ||||||:|||||||||||||||||:| |||||||||  :|||| ||||||||||||||||
Db        342 TTAAATATCAGGCTCTATTACTTGAAGGTCCAGTGCTGTGACTGCGCACTTGTGCAACTC 401

Qy       3776 TTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAAR 3835
             |||||||||:| |||||||||||||||||||||||||||:||: |||||||||||||:
Db        402 TTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAATGTAACTGTCAACAAG 461

Qy       3836 TAAT 3839
              ||||
Db        462 TAAT 465


RESULT 194
AQ221075/c
LOCUS       AQ221075                 431 bp    DNA     linear   GSS 19-SEP-1998
DEFINITION  HS_2004_B2_H12_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2004 Col=24 Row=P, genomic survey
            sequence.
ACCESSION   AQ221075
VERSION     AQ221075.1  GI:3634688
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  PUBMED    10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2004  row: P  column: 24
            Class: BAC ends
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2004 Col=24 Row=P"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.0%;  Score 374.4;  DB 1;  Length 431;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  386;  Conservative   39;  Mismatches    6;  Indels    4;  Gaps    3;

Qy       2981 TGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCA 3040
             ||||||||:|||:|||||||||||||| ||:||:|||::|||||||||||||||||||||
Db        431 TGTCCTTCAGTATGTGGATGATTTACTATTAGCTGCCTGTTCAGAAACCTTGTGCCATCA 372

Qy       3041 AGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARA 3100
             |||||||||||:|||||:||||||||||:|||||||||||:|||||||||||   :|
Db        371 AGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAAGAGAA 312

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
             :|||||:|||||||||||||||||||||||||||||:||||:||||||||||||||||:
Db        311 GGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAAATTATCCAAAGGCACCAA 252

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
             ||||||||||||||:::|||||||||||||||||||||||||||||:||||||||||||
Db        251 GGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCTCAAAACCCTAAA 192

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
             |||||||||:|::|||||||||:|||:|||:|||||||||||:|||||||  |||||||:||
Db        191 GCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT-CCCAGGTATG 133
```

```
Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:||||||||||:|||||:|||||:||||||||||||||||||||||||||||||||||
Db        132 GCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAAGCCAATACCCATT 73

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:|||||||||:|:||||||:::|||||||||| ||||  |||||||||||||| |||
Db          72 TAGTAAGATGGATACCTGAAGCAAAAGTGGCTTTACAGG--CCTAAAGAAGGCCTT-AAC 16

Qy       3401 CCAAGYCCCAGTGTT 3415
              ||||:|||||||||
Db          15 NCAAGTCCCAGTGTT 1
```

RESULT 195
BI056595/c

| | |
|---|---|
| LOCUS | BI056595 376 bp mRNA linear EST 15-JUN-2001 |
| DEFINITION | RC5-GN0282-210201-013-G04 GN0282 Homo sapiens cDNA, mRNA sequence. |
| ACCESSION | BI056595 |
| VERSION | BI056595.1 GI:14464125 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 376) |
| AUTHORS | Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J. |
| TITLE | Shotgun sequencing of the human transcriptome with ORF expressed sequence tags |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) |
| PUBMED | 10737800 |
| COMMENT | Contact: Simpson A.J.G.<br>Laboratory of Cancer Genetics<br>Ludwig Institute for Cancer Research<br>Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil<br>Tel: +55-11-2704922<br>Fax: +55-11-2707001<br>Email: asimpson@ludwig.org.br<br>This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-210201-013-G04&t3=2001-02-21&t4=1)<br>Seq primer: puc 18 forward<br>High quality sequence stop: 376. |

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .376<br>/organism="Homo sapiens"<br>/mol_type="mRNA"<br>/db_xref="taxon:9606"<br>/dev_stage="Adult"<br>/clone_lib="GN0282"<br>/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions." |

```
  Query Match             5.0%;  Score 374.2;  DB 1;  Length 376;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches  373;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 317

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 257

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        256 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 197

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 137

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        136 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 77

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 17

Qy       6358 CAAATAGTCTGCCTA 6372
              |||||||||||||||
Db          16 CAAATAGTCTGCCTA 2
```

RESULT 196
AQ536209

| | |
|---|---|
| LOCUS | AQ536209 521 bp DNA linear GSS 18-MAY-1999 |
| DEFINITION | RPCI-11-321K3.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-321K3, genomic survey sequence. |
| ACCESSION | AQ536209 |
| VERSION | AQ536209.1 GI:4847899 |
| KEYWORDS | GSS. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 521) |
| AUTHORS | Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C. |
| TITLE | Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building |
| JOURNAL | Unpublished (1997) |
| COMMENT | Other_GSSs: RPCI-11-321K3.TV<br>Contact: Shaying Zhao, William Nierman, Mark Adams<br>Department of Eukaryotic Genomics<br>The Institute for Genomic Research<br>9712 Medical Center Dr., Rockville, MD 20850<br>Tel: 301 838 0200<br>Fax: 301 838 0208<br>Email: hbe@tigr.org<br>Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.<br>Seq primer: SP6<br>Class: BAC ends. |

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .521<br>/organism="Homo sapiens"<br>/mol_type="genomic DNA"<br>/db_xref="GDB:7623122"<br>/db_xref="taxon:9606"<br>/clone="RPCI-11-321K3"<br>/sex="Male"<br>/cell_type="Lymphocytes"<br>/clone_lib="RPCI-11"<br>/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library" |

```
  Query Match             5.0%;  Score 373.3;  DB 1;  Length 521;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  437;  Conservative   19;  Mismatches   52;  Indels    9;  Gaps    4;

Qy       4147 GGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATT 4206
             ||||||||||||||||:|||:||| ||||||||:|||||| ||||||||||||||:||||||
Db          7 GGAATTCCTAACTTCTGAGGGAATACCTATCAAACATCAAGAAGCCATTAGGAGATTATT 66

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGG 4266
             |:||||:|||||||||||||:|||||||  |||||||||||||: |||||||| |||||
Db         67 ATTGGCTGTACAGAAACCTAAAGAGGTGATAGTCTTACACTGCCAGGGTCATCAAAAAGG 126

Qy       4267 AAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAG 4326
             |||| ||||| |||||:|||:|||: ||||||| :||||||||| :|||| |||  || |
Db        127 AAAG-AAAGGAAAATAGAAGGGAACCGCCAAGCGGATATTGAAGTCAAAAAAGCCACATG 185

Qy       4327 GCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCG 4386
             |||||||||| |||||||||||||||||  ||   ||| |||||||||||||||||| ||||
Db        186 GCAGGACCCTTCATTAGAAATGCTTATAGAAAGGCCCCTAGTATAGGGTAATCCCCTCCG 245

Qy       4387 GGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAG 4443
             ||||||||  |||||||||||||||||||| ||| ||||| ||||||||||||||     |
Db        246 GGAAACCATACCCCAGTACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGAACAAAC 305

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||  |||||||   || || |||||||  ||||||| |||||||||||| |||| |||
Db        306 TTCCCTCCCCTTCAGATGGCTAGCCACCAAAGAAGGAAAAATACTTTTGCTTGCAGCTAA 365

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA 4563
             | ||| ||||||||||||||||||||| |||||||| || |||||||| ||||||||||
Db        366 CTAATAGAAATTACTTAAAACCCTTCACCAAACCTTCCA-TTAGGCATTGATAGCACCCA 424

Qy       4564 TCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC----ARA 4619
             |||:|||||||||| |||||||||| |||||||| ||||||||||||||||    |:|
Db        425 TCAGATGGCCAAATTATTATTTACTGGTCCAGGCCTATTCAAAACTATCAAGCAGATAGA 484

Qy       4620 TAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCC 4656
             ||:||| ||||||||||:|||||||:|:|||||||||
Db        485 TAGTCAAGGCCTGTGAAGTGTGCCAAAGAAATAATCC 521


RESULT 197
AQ167811
LOCUS       AQ167811                434 bp    DNA     linear   GSS 16-OCT-1998
DEFINITION  HS_2203_B2_G01_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2203 Col=2 Row=N, genomic survey
            sequence.
ACCESSION   AQ167811
VERSION     AQ167811.1  GI:3561526
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 434)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2203  row: N  column: 2
            Class: BAC ends
            High quality sequence stop: 434.
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2203 Col=2 Row=N"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             5.0%;  Score 372.7;  DB 1;  Length 434;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  382;  Conservative   25;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       3515 CYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAY 3574
             |::||||::|||||| |||||| |:||||||||||||||||| |||:|:|||||| |||:
Db          2 CTATGGCATACCTGAATAAGGATATTGATGTAGTGGCAAAGGTTTGGCCTCATTGATTAT 61

Qy       3575 GGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAG 3634
             ||||||||   ||||||||||::|||||||||||||||||||||||||||||||:||||
Db         62 GGGTAGTGACAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAG 121

Qy       3635 ATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGT 3694
             |||||||||||| |||||||  ||||:|||||::||| |||||||||||||||||||||||
Db        122 ATCTTACTGTGTAGACATCTAATGATGTGAACGGCACACTCACTGCTAAAGGAGACTTGT 181

Qy       3695 GGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGC 3754
             |||||||||||||:||||||||||||:|||||||||||||||||||:| |||||||||
Db        182 GGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCAGTGCTGT 241

Qy       3755 RACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGA 3814
             :|||| |||||||||| ||||||||||:| |||||||| ||||||||||||||||||||
Db        242 GACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAATGAAGAAAAGA 301

Qy       3815 TARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAG 3874
             ||:||:|||||||||||:||||| |||||||||||||| ||| | |||||||| |||
Db        302 TAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAAGGGACCTTNTAG 361

Qy       3875 ARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAA 3934
             |:|||||:|||||||||||: ||| |||||| ||||||||||||||||||||||||||||
Db        362 AGGTTCCCTTGACTGATCCCAACC-TCAACTCGTATACTGATGGAAGTTCCTTTGTAGAA 420

Qy       3935 AAAGGACTTCGAA 3947
             ||| |||||| ||
Db        421 AAAAGACTTCAAA 433


RESULT 198
BF993049
LOCUS       BF993049                379 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-311000-204-d04 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF993049
VERSION     BF993049.1  GI:12399372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
```

```
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            311000-204-d04&t3=2000-10-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 323.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.0%;  Score 372.4;  DB 1;  Length 379;
  Best Local Similarity   96.0%;  Pred. No. 0;
  Matches  364;  Conservative   14;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4479 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 4538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 60

Qy       4539 TTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCC 4598
              |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db         61 TTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCC 120

Qy       4599 TTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCC 4658
              ||||||||||||||||||:|||:|||||||||||||:|||||||:|:||||||||||||
Db        121 TTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCC 180

Qy       4659 TGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACT 4718
              |||||:|||||||||||||||||:||||||:|||||||||||||||||:|:||:|||||
Db        181 TGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACT 240

Qy       4719 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 4778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 300

Qy       4779 ARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGT 4838
              |:||||||||||||||||||:|||||||| ||||||||||||||||||||||||||||||
Db        301 AGATACTTTCACGGGTTGGGCAGAGGCCTTCCNCTGTAGGACAGAAAAGGCCCAAGAGGT 360

Qy       4839 AATAAAGGCACTAGTTCAT 4857
              ||||||||||||||||||
Db        361 AATAAAGGCACTAGTTCAT 379


RESULT 199
BI052559/c
LOCUS       BI052559                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-A04 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052559
VERSION     BI052559.1  GI:14460089
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-A04&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.0%;  Score 372.2;  DB 1;  Length 373;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches  371;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          73 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
              |||||||||||||
Db          13 CAAATAGTCTGCC 1
```

```
RESULT 200
BI052567/c
LOCUS       BI052567                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-E06 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052567
VERSION     BI052567.1  GI:14460097
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-E06&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.0%;  Score 372.2;  DB 1;  Length 373;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches  371;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          73 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
              |||||||||||||
Db          13 CAAATAGTCTGCC 1


RESULT 201
BI056593/c
LOCUS       BI056593                374 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-210201-013-F06 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056593
VERSION     BI056593.1  GI:14464123
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            210201-013-F06&t3=2001-02-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 364.
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             5.0%;  Score 371.6;  DB 1;  Length 374;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches  371;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 315
```

```
Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 255

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTC 6176
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        254 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTC 195

Qy       6177 AATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTT 6236
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        194 AATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGCTTT 135

Qy       6237 AGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAA 6296
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        134 AGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAA 75

Qy       6297 ATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAC 6356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          74 ATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAC 15

Qy       6357 ACAAATAGTCTGCC 6370
              ||||||||||||||
Db          14 ACAAATAGTCTGCC 1
```

```
RESULT 202
AQ734945/c
LOCUS       AQ734945                449 bp    DNA     linear   GSS 15-JUL-1999
DEFINITION  HS_3085_B2_D07_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3085 Col=14 Row=H, genomic survey
            sequence.
ACCESSION   AQ734945
VERSION     AQ734945.1  GI:5506497
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3085  row: H  column: 14
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3085 Col=14 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.9%;  Score 370.7;  DB 1;  Length 449;
  Best Local Similarity   88.2%;  Pred. No. 0;
```

```
  Matches  389;  Conservative   26;  Mismatches   21;  Indels    5;  Gaps    2;

Qy       3713 TACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAA 3772
              | | |||||:|||| || ||||||||||:||||||||||||:||||||||||||||||||
Db        448 TTCATAAATATCAGACTGTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAA 389

Qy       3773 CTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAAC 3832
              ||||||||||||:| ||| ||||||||||||||||||||||:||:||||||||||||||
Db        388 CTCTTAACCCAGCCACATGTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAAC 329

Qy       3833 AARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATC 3892
              ||:||||| ||||| | ||||||||| ||||||||||:||| :|||||:||||||||||
Db        328 AAGTAATTGCTCAACCTTATGCCACTTGAGGGGACCTTTTAGGGGTTCCCTTGACTGATC 269

Qy       3893 CYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYG 3952
              |: ||| |||||||||||||||| |||||||||||||||||| ||||||| ||| |:|
Db        268 CCAACC-TCAACTTGTATACTGATGAAAGTTCCTTTGTAGAAAACGGACTTCAAAATGCG 210

Qy       3953 GGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACT 4012
              ||||||||     ||||||||||||||||:|||||||||||||||||||||||||| |
Db        209 GGGTATGC----CTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAAAT 154

Qy       4013 AGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAA 4072
              ||||||:||||:|||||||||||||||:|||:| |||||||||||||||||||:||||
Db        153 AGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGGAAAA 94

Qy       4073 AGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCA 4132
              ||||:|||||||||||||||:|:||||||:|||||||| ||||||||||::|||
Db          93 AGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCA 34

Qy       4133 ATATGSARAGAAAGGGAATTC 4153
              |||||:|:||||| |||||||
Db          33 ATATGGAGAGAAAAGGAATTC 13
```

```
RESULT 203
BI052561/c
LOCUS       BI052561                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-B03 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052561
VERSION     BI052561.1  GI:14460091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-B03&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
```

```
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.9%;  Score 369;  DB 1;  Length 373;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  369;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          73 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAGCTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
             ||||||||| |||
Db          13 CAAATAGTCAGCC 1


RESULT 204
BI052571/c
LOCUS       BI052571                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-G11 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052571
VERSION     BI052571.1  GI:14460101
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-G11&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 373.
FEATURES            Location/Qualifiers
     source         1. .373
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="GN0282"
                    /note="Organ: placenta_normal; Vector: puc18; Site_1:
                    SmaI; Site_2: SmaI; A mini-library was made by cloning
                    products derived from  ORESTES PCR (U.S. Letters Patent
                    application No. 196,716 -  Ludwig Institute for Cancer
                    Research) profiles into  the pUC 18 vector. Reverse
                    transcription of tissue mRNA and cDNA amplification were
                    performed under low stringency conditions."

  Query Match             4.9%;  Score 369;  DB 1;  Length 373;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  369;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCCCAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          73 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
             ||||||||| |||
Db          13 CAAATAGTCAGCC 1


RESULT 205
BI052572/c
LOCUS       BI052572                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-H10 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052572
VERSION     BI052572.1  GI:14460102
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
```

```
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-H10&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.9%;  Score 369;  DB 1;  Length 373;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  369;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCCCAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          73 TTTAGCAACACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
              |||||||||||||
Db          13 CAAATAGTCTGCC 1


RESULT 206
BI056589/c
LOCUS       BI056589                373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-210201-013-C03 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056589
VERSION     BI056589.1  GI:14464119
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            210201-013-C03&t3=2001-02-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 344.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.9%;  Score 369;  DB 1;  Length 373;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  369;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAGTGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74
```

```
Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACA 6357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
             ||||||||| |||
Db         13 CAAATAGTCAGCC 1


RESULT 207
AQ064818
LOCUS       AQ064818               467 bp    DNA     linear   GSS 04-AUG-1998
DEFINITION  HS_2221_A2_F03_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2221 Col=6 Row=K, genomic survey
            sequence.
ACCESSION   AQ064818
VERSION     AQ064818.1  GI:3380646
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2221  row: K  column: 6
            Class: BAC ends
            High quality sequence stop: 467.
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2221 Col=6 Row=K"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.9%;  Score 368.5;  DB 1;  Length 467;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  391;  Conservative   23;  Mismatches   48;  Indels    1;  Gaps    1;

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||: || | | || ||||||| || |:||  |  |||:||||||| | ||||||||||
Db          1 GAAATAGANGAATTGTCAAAGGGNTGTCCTCTATTGTTACGGGTAGTAGCGGCAGTAGCA 60

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|| |||||||||||||||||||||||||||:|||||| ||| ||||||||||||||
Db         61 GTCTTATTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATTTTATTGTGTGGACATCT 120

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             ||||:|||||::||||||| |||||||||||||||||||||||||||||||||:|||||
Db        121 TATGATGTGAATGGCATACTGACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 180

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:||| |||||||||||||||:||||||||||  :||||.  ||||||||||||
Db        181 CTTAAATATCACGCTCTATTACTTGAAGGGCCAGTGCTGTGACTGCATACTTGTGCAACT 240

Qy       3775 CTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAA 3834
             ||||| ||||:  |||||||||||||||| ||||||||||:||:|||||||||||||||
Db        241 CTTAATCCAGCAACATTTCTTCCAGACAATAAAGAAAAGATAGAACATAACTGTCAACAA 300

Qy       3835 RTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCY 3894
             :|||| |||| ||||| || |||||| ||||||||:||||:|||||:||||||||||:
Db        301 GTAATGGCTCANACCTACGCTGCTCGATGGGACCTTTTAGAGGTTCCCTTGACTGATCCC 360

Qy       3895 GACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGG 3954
              ||| ||| |||||||||||||||||||||||||| ||||| || |||| | || |||:|
Db        361 AACC-TCAGCTTGTATACTGATGGAAGTTCCTTNGTAGATAATGGACCTTGATAAGCGAT 419

Qy       3955 GTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
             |||||||||||||||||||| ||||:|| ||||||  |||
Db        420 GTATGCAGTGGTCAGTGATAATGAAATACTTAAAAGTATCCCC 462


RESULT 208
AQ057826
LOCUS       AQ057826               443 bp    DNA     linear   GSS 30-JUL-1998
DEFINITION  CIT-HSP-2339F8.TR CIT-HSP Homo sapiens genomic clone 2339F8,
            genomic survey sequence.
ACCESSION   AQ057826
VERSION     AQ057826.1  GI:3354352
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2339F8.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2339F8"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.9%;  Score 368.1;  DB 1;  Length 443;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  369;  Conservative   34;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
             | |||:||||||||:|||||||| ||||:||:|||::|||||| |||||||||:||:|||
Db          22 TCAAGCTTGCCAACGGGGCAAGAGTTTTCTTTATATGTCACAG-AAAAAACAGGAATAGC 80

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
             |||:| ||||||||||||:|| :|||||:||||||||||:: |||::||||||:|||
```

```
Db         81 TCTAGAAGTCCTTACACAGGTCTGAGGGATGAGCTTGCAACCCACGGCATACCTGAGTAA 140

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
              || ||: |||||| |||||||||||||:|:||||||||:||||||| |||||||||
Db         141 GGCAATCGATGTAATGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGC 200

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATC 3653
              |||::|||||||||||||||||||||||||||||:|||||||||||||||||||| ||
Db         201 AGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACTTC 260

Qy       3654 TCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTT 3713
              ||||||:|||||: |||||||||||||||||||||||||||||||||||||||||: |||
Db         261 TCATGATGTGAACTGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTT 320

Qy       3714 ACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAAC 3773
              | ||||||:|||||||||||||||||:||| ||||||| :|||| ||||||||||||
Db         321 ATTTAAATATCAGGCTCTATTACTTGAAGGGCTAGTGCTGTGACTGCACACTTGTGCAAC 380

Qy       3774 TCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACA 3833
              ||||||||||:| |||||||||||||||||||| |||||:||:||||||||||||||
Db         381 TCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAACAGATAGAACATAACTGTCAACA 440

Qy       3834 ART 3836
              |:|
Db         441 AGT 443


RESULT 209
AQ034878
LOCUS       AQ034878               423 bp    DNA     linear   GSS 11-JUL-1998
DEFINITION  CIT-HSP-2316B5.TR CIT-HSP Homo sapiens genomic clone 2316B5,
            genomic survey sequence.
ACCESSION   AQ034878
VERSION     AQ034878.1  GI:3301052
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2316B5"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.9%;  Score 367.9;  DB 1;  Length 423;
  Best Local Similarity   88.7%;  Pred. No. 0;

  Matches  376;  Conservative   25;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:||||||||||||||| |||:|:|||||| |||:||
Db           1 ATGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGTTTGGCCTCATTGATTATGG 60

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||||||   ||||||||||::|||||||||||||||||||||||||||:||||||
Db          61 GTAGTGACAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 120

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              |||||||||| ||||||| ||||:|||||::||| |||||||||||||||||||||||
Db          121 CTTACTGTGTAGACATCTAATGATGTGAACGGCACACTCACTGCTAAAGGAGACTTGTGG 180

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||:|||||||||:|||||||||||||||||:|  |||||||| :|
Db          181 CTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTTCAGTGCTGTGA 240

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATA 3816
              ||| |||||||||| ||||||||||:| |||||||| |||||||||||||||||||||
Db          241 CTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAATGAAGAAAAGATA 300

Qy       3817 RAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGAR 3876
              :||:|||||||||||:||||| |||||||||||||| ||| ||||||||||:||||:
Db          301 GAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGGGACCTTTTAGAG 360

Qy       3877 GTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 3936
              |||||:||||||||||: ||| |||||| |||||||||||||||||| |||||||||
Db          361 GTTCCCTTGACTGATCCCAACC-TCAACTAGTATACTGATGGAAGTTCCTATGTAGAAAA 419

Qy       3937 AGGA 3940
              ||||
Db          420 AGGA 423


RESULT 210
BI052570/c
LOCUS       BI052570               373 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-G07 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052570
VERSION     BI052570.1  GI:14460100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-G07&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
```

```
    source          1. .373
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="GN0282"
                    /note="Organ: placenta_normal; Vector: puc18; Site_1:
                    SmaI; Site_2: SmaI; A mini-library was made by cloning
                    products derived from  ORESTES PCR (U.S. Letters Patent
                    application No. 196,716 -  Ludwig Institute for Cancer
                    Research) profiles into  the pUC 18 vector. Reverse
                    transcription of tissue mRNA and cDNA amplification were
                    performed under low stringency conditions."

  Query Match             4.9%;  Score 367.4;  DB 1;  Length 373;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  368;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5998 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 6057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGT 314

Qy       6058 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 6117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCG 254

Qy       6118 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCA 6177
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        253 GCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCA 194

Qy       6178 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 6237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTA 134

Qy       6238 GTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 6297
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        133 GTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAA 74

Qy       6298 TTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACA 6357
             ||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          73 TTTAGCAATCCTACATACACAACCAACTCCCAATGCATCAGGCGGGTAACTCCTCCCACA 14

Qy       6358 CAAATAGTCTGCC 6370
             |||||||||||| |
Db          13 CAAATAGTCTGTC 1


RESULT 211
BI052566/c
LOCUS       BI052566                379 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-220101-011-E05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI052566
VERSION     BI052566.1  GI:14460096
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            220101-011-E05&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 363.
FEATURES            Location/Qualifiers
     source         1. .379
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="GN0282"
                    /note="Organ: placenta_normal; Vector: puc18; Site_1:
                    SmaI; Site_2: SmaI; A mini-library was made by cloning
                    products derived from  ORESTES PCR (U.S. Letters Patent
                    application No. 196,716 -  Ludwig Institute for Cancer
                    Research) profiles into  the pUC 18 vector. Reverse
                    transcription of tissue mRNA and cDNA amplification were
                    performed under low stringency conditions."

  Query Match             4.9%;  Score 367.2;  DB 1;  Length 379;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches  366;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6003 CACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCA 6062
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCA 309

Qy       6063 TACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCA 6122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCA 249

Qy       6123 AAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCC 6182
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        248 AAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCC 189

Qy       6183 TGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGG 6242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGG 129

Qy       6243 ACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAG 6302
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        128 ACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAG 69

Qy       6303 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 6362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 CAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAAT 9

Qy       6363 AGTCTGCC 6370
             ||||||||
Db          8 AGTCTGCC 1


RESULT 212
AQ216381
LOCUS       AQ216381                408 bp    DNA     linear   GSS 19-SEP-1998
DEFINITION  HS_2189_A1_F07_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2189 Col=13 Row=K, genomic survey
            sequence.
ACCESSION   AQ216381
VERSION     AQ216381.1  GI:3631701
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2189  row: K  column: 13
            Class: BAC ends
            High quality sequence stop: 408.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2189 Col=13 Row=K"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.9%;  Score 366.1;  DB 1;  Length 408;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  364;  Conservative   32;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       3431 GCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACAC 3490
             ||||||||||:||:|||::||| | |||||| |||:||:||||||:|||||||||||||
Db             1 GCAAGACTTTTCTTTATATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACAC 59

Qy       3491 AGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGG 3550
             ||:|||:|||||:|||||||||||||::||||::||||| ||||||||:||||||||||
Db            60 AGGTCCGAGGGACGAGCTTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGG 119

Qy       3551 CAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAG 3610
             ||||||||||:|:|||||||||:|||||||| ||||||||||::||||||||||||||
Db           120 CAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAG 179

Qy       3611 CAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCA 3670
             ||||| ||||||||||||:|||||| |||||||||||||||||||:||||||::|||
Db           180 CAGTTTAAATAATACAGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCA 239

Qy       3671 TACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTC 3730
             |||| |||||||||||||||||||||||||||||:||||||||||||:||||||||
Db           240 TACTGACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTC 299

Qy       3731 TATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCAT 3790
             ||||||||||:|||||||||||:|||| ||| |||||||||||||||||||:| |||
Db           300 TATTACTTGAAGGGCCAGTGCTGCGACTGCGCATTTGTGCAACTCTTAACCCAGTCCCAT 359

Qy       3791 TTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAAT 3839
             ||||||| ||||||||||||||||:||:|||||||||||||:||||
Db           360 TTCTTCCGGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAAT 408


RESULT 213
AA837267/c
LOCUS       AA837267               443 bp    mRNA    linear   EST 31-MAR-1998
DEFINITION  od26b10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1369051
            similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AA837267
VERSION     AA837267.1  GI:2912466
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1375   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1369051"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from human tonsillar cells
                     enriched for germinal center B cells by flow sorting
                     (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
                     David Allman (NCI) and Dr. Gerald Marti (CBER).  cDNA
                     synthesis was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

  Query Match             4.9%;  Score 365.5;  DB 1;  Length 443;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   34;  Indels    1;  Gaps    1;

Qy       7147 ATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGT 7206
             |||||| |||||||||| |||||| |||||| |||||||||||||||||||| |||: |
Db            443 ATCTCAACTGCACAACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCAT 384

Qy       7207 CGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGG 7266
             | ||||||||| |||||||||||| |||||||||||||||  ||||||||||||||||
Db            383 CAGCCAACCTCACCAACAGCACTTGGGTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGG 324

Qy       7267 ACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACA 7326
             |||||||||||||||||| ||:|||||||||:||| |||||:|||||||||||||| ||
Db            323 ACTAGCTGGATTTCCTAGGCCGATTAAGAATCCCTAAACCTAGCTGGGAAGGTGACCGCA 264

Qy       7327 TCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACT 7386
             |||||||||||||||||||||||||||||||||:|||||||  ||||||||||||||||||
Db            263 TCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACT 204

Qy       7387 AAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGA 7446
             |||||||||||||||||| |||||||||||||||||||||||||:||||||:|||||
Db            203 AAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGA 144

Qy       7447 GCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAA 7506
             ||||||  |||||||||| ||| |  |||||||||| |  |||||||| || |  ||||
Db            143 GCACAGTGGGAGGGACAAGGATTGCAATATAAACCCAGGCATTCGAGCCAGC-ANAGCAA 85

Qy       7507 CCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATT 7566
```

```
                 || |||||||||||| |||| ||||||||||||||||||||  ||||||||||||||||
Db         84 CCGCCTTTGGGTCCCTTCCCCTTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTATT 25

Qy       7567 AAATCTTGCARCTG 7580
                ||||||||||:|||
Db         24 AAATCTTGCAACTG 11


RESULT 214
H01325/c
LOCUS       H01325                  388 bp    mRNA    linear   EST 19-JUN-1995
DEFINITION  yi99e01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:147384 3', mRNA sequence.
ACCESSION   H01325
VERSION     H01325.1  GI:864258
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 790
            High qality sequence stops: 346
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 790   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:559031"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:147384"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.8%;  Score 363.2;  DB 1;  Length 388;
  Best Local Similarity   95.3%;  Pred. No. 0;
  Matches  362;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       7198 AGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTG 7257
                | ||||:|||||| |||||||||||| ||||||||||||||||||||||||||||||||
Db        385 AACGGTCGTCGGCNAACCTCCCCAANAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTG 326

Qy       7258 AGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAG 7317
                ||||||||||||||||||||||||||||||:||||||||:|||||||||:|||||||
Db        325 AGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAG 266

Qy       7318 GTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGA 7377
                ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        265 GTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGA 206

Qy       7378 GAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTAT 7437
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        205 GAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTAT 146

Qy       7438 TGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGG 7497
                |||:||||||||||||||||||||||||||||||||||||||||||||:|||||  |||
Db        145 TGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAAGCGG 86

Qy       7498 CAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTT 7557
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 CAACGGCAAGCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTT 26

Qy       7558 CACTCTATTAAATCTTGCAR 7577
                |||||||||||||| |  |:
Db          25 CACTCTATTAAATCCTTGAA 6


RESULT 215
AI074704/c
LOCUS       AI074704                470 bp    mRNA    linear   EST 06-AUG-1998
DEFINITION  ox83d05.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1662921
            3' similar to contains PTR7.t2 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AI074704
VERSION     AI074704.1  GI:3401348
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1662921"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-PacI;
                     Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid
                     DNA from three normalized libraries (melanocyte 2NbHM,
                     pregnant uterus NbHPU, and fetal heart NbHH19W) were
                     mixed, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from pools of 5,000 clones made from the same 3 libraries.
                     The pools consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."

  Query Match             4.8%;  Score 363.2;  DB 1;  Length 470;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  405;  Conservative    7;  Mismatches   35;  Indels   10;  Gaps    2;
```

```
Qy       7109 CTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTAC 7168
             | |||||| |||||||||||||||||||||||||||||| ||||| |||||| ||| ||||
Db        470 CCGATGTTGATGACATCAAAGGCACCCCTCCTGAGGAACTCTCAACTGCACGACCCCTAC 411

Qy       7169 TACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCAC 7228
             || |||||||||||||||||||||||||||||| |||: | ||||||||||||||||||||
Db        410 TATGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTCATTGGCCAACCTCCCCAACAGCAC 351

Qy       7229 TTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 7288
             || |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db        350 TTGGGTTTTCCTGTTGAGAGTGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 291

Qy       7289 AYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 7348
             |:|||||||||:|||||||||:|||||||||||||  |||||||||||||||||||||||||
Db        290 ACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACGGCATCCACCTTTAAACACGGGGCTT 231

Qy       7349 GCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTA 7399
             |||||||||:|||||| ||||||||          ||| |||||||||||||||||||||
Db        230 GCAACTTAGCTCACACCCGACCAATCAGGGAGTAAAGACAGCTCACTAAAATGCTAATTA 171

Qy       7400 GGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGG 7459
             ||| || |||||||||||||||||||||||||||||:||| ||:||||||||| | ||||||
Db        170 GGCTAAAACAGGAGGTAAAGAAATAGCCAATCATCTATCGCCTGAGAGCAC-GAGGGAGG 112

Qy       7460 GACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTC 7519
             |||||||||| |||||||||||||| |  ||| ||||||||  ||| |||| |||||||||
Db        111 GACAATGATCAGGATATAAACCCAGGCATTCAAGCCGGCAGTGGCTACCCGCTTTGGGTC 52

Qy       7520 CCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATT 7556
             |||||  |||||||||||||||||||||||  |||||
Db          51 CCCTCTGTTTGTATGGGAGCTCTGTTTTCACTCTATT 15
```

RESULT 216
R77278/c
LOCUS R77278 438 bp mRNA linear EST 06-JUN-1995
DEFINITION yi75d06.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145067 3', mRNA sequence.
ACCESSION R77278
VERSION R77278.1 GI:851910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 438)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1163
High quality sequence stops: 273
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1163 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 273.
FEATURES Location/Qualifiers
source 1. .438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:560657"
/db_xref="taxon:9606"
/clone="IMAGE:145067"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta Nb2HP"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. "

Query Match 4.8%; Score 360.8; DB 1; Length 438;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 404; Conservative 8; Mismatches 22; Indels 4; Gaps 4;

```
Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
             | |||| ||||||||| |||||| |||||||||| |||||||||| || |||||| |||
Db        438 CCCCTGGGGAAATCTCGGCTGCAAAACCTCTACTAAGCCCCAATTCGGCGGGAAGCNGTT 379

Qy       7196 AG-AGCGGTSGTCGGCCAACCT-CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGG 7253
             || | ||||:|| |||||||||| |||||| |||||||||||  ||||||||||||||||
Db        378 AGAAACGGTCGTNGGCCAACCTCCCCCAANAGCACTTAGGTTNCCCTGTTGAGATGGGGG 319

Qy       7254 ACTGAGAGACAGGACTAGCTGGATTT-CCTAGGCTGA-YTAAGAATCCYTAAGCCTAGST 7311
             |||||||  |||||||||||||||||| |||||||||| :|||||||||:|||||||||:|
Db        318 ACTGAGANACAGGACTAGCTGGATTTCCCTAGGCTGACCTAAGAATCCCTAAGCCTAGCT 259

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
             ||||||||| |||||||||||||||||||||||||||||||||||:|||||||||||||
Db        258 GGGAAGGTGNCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCA 199

Qy       7372 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 7431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATC 139

Qy       7432 ATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCG 7491
             ||:||||||:|||||||||||||||||||||||||| |||||||||||||||||||:||||
Db        138 ATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGANCGGGATATAAACCCAAGTCTTCG 79

Qy       7492 AGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATG 7551
                |||||||||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db          78 NNNCGGCAACGGCAATCCCCTTTGGGTCCCCTCCCNTTGTATGGGAGCTCTGTTTTCATG 19

Qy       7552 CTATTTCACTCTATTAAA 7569
             ||||||| ||||||||||
Db          18 CTATTTCNCTCTATTAAA 1
```

RESULT 217
CB069106/c
LOCUS CB069106 446 bp mRNA linear EST 21-JAN-2003
DEFINITION is13a01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6364680 3', mRNA sequence.
ACCESSION CB069106
VERSION CB069106.1 GI:27813626
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 446)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,

```
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 419.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6364680"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

  Query Match             4.8%;  Score 359.5;  DB 1;  Length 446;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  401;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7133 CCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCA 7192
             ||| ||  |||||||||||| |||||||||| ||||||| ||||||||||||||||||||
Db        446 CCCGTCTCGAGGAAATCTCAACTGCACAACCCCTACTACACCCCAATTCAGCAGGAAGCA 387

Qy       7193 GTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGG 7252
             ||||||||||:||| |||||||||||||| |||||||| ||||||||||||||||  |||
Db        386 GTTAGAGCGGTCGTCAGCCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGCGGG 327

Qy       7253 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTG 7312
             ||||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||:||
Db        326 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTG 267

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
             |||||||||||  | || |||||||||||||||||||||||||:|||||||  |||||
Db        266 GGAAGGTGACCGCGTCTACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCAACCAA 207

Qy       7373 TC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAAT 7423
              ||         |||||||||||||||||||||||||||||| ||||||||| ||||||
Db        206 TCAGATAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAGAGAAAT 147

Qy       7424 AGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 7483
             ||||||||||:||| ||:|||||||||||||||||||||||||| ||||||||||||||
Db        146 AGCCAATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCCGGATATAAACCCA 87

Qy       7484 AGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTG 7543
              ||.|||||||| |||| ||| |||| ||||| ||||||||||||||||||||||||||
Db         86 AGCATTCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTGTATGGGAGCTCTA 27

Qy       7544 TTTTCATGCTATTTCACTCTATTAAATCTTGCARCT 7579
              |          |||||||||||||||||||||:||
Db         26 T----------TTTCACTCTATTAAATCTTGCAACT 1
```

```
RESULT 218
DB374307/c
LOCUS       DB374307                399 bp    mRNA    linear   EST 11-DEC-2005
DEFINITION  DB374307 PLACE1 Homo sapiens cDNA clone PLACE1003157 3', mRNA
            sequence.
ACCESSION   DB374307
VERSION     DB374307.1  GI:83566911
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1003157"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             4.8%;  Score 358.7;  DB 1;  Length 399;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  371;  Conservative    7;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       7182 AGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTG 7241
             |||||||||||||||||||  |  | ||| ||| |||| ||||| | ||||||||||||||
Db        399 AGCAGGAAGCAGTTAGAGCGTTTNTGGGCAAACTTCCCNAACAGAATTTAGGTTTTCCTG 340

Qy       7242 TTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYT 7301
             ||||||  ||||||  |||||||||||||||| ||||||| ||||||||:|||||||:|
Db        339 TTGAGANGGGGGANTGAGAGACAGGACTAGNTGGATTTCTTAGGCTGATTAAGAATCCCT 280

Qy       7302 AAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCA 7361
             ||||||||  |||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        279 AAGCCTAGNTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCA 220

Qy       7362 CACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAA 7421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAA 160

Qy       7422 ATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACC 7481
             |||||||| ||:|||||:||||||||||||||||||||||||||||||||||||||||||
Db        159 ATAGCCAATNATTTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACC 100

Qy       7482 CAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTC 7541
             ||| |:|| |||||||||||||||||||||||| |||||| |||||||||||||||| ||
Db         99 CAANTTTTNGAGCCGGCAACGGCAACCCCCTTT-GGTCCCNTCCCTTTGTATGGGAGTTC 41
```

```
Qy       7542 TGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||||| | ||||||||||||||:||||
Db         40 TGTTTTCATGCTATTTCATTNTATTAAATCTTGCAACTGC 1


RESULT 219
AQ701392/c
LOCUS       AQ701392                442 bp    DNA     linear   GSS 07-JUL-1999
DEFINITION  HS_2133_A2_F11_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2133 Col=22 Row=K, genomic survey
            sequence.
ACCESSION   AQ701392
VERSION     AQ701392.1  GI:5410818
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2133  row: K  column: 22
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2133 Col=22 Row=K"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.8%;  Score 358;  DB 1;  Length 442;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  370;  Conservative   41;  Mismatches   27;  Indels    2;  Gaps    2;

Qy       2947 AAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTAC 3006
              ||||| ||||:||:|::|||||||||||||||||||||||||:|||   ||||||||||
Db        442 AAGACGTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCGGTACAGGGATGATTTAA 383

Qy       3007 TTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCC 3066
              ||||:||: ||::|||||||| ||| |||| |||||||||||||| :|||||:| ||||
Db        382 TTTTAGCCACCTATTCAGAAATCTTATGCCGTCAAGCCACCCAAGTGCTCTTAAGTTTCT 323

Qy       3067 TCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTA 3126
              || |:||||||||||||| ||||||||||:|:|:|||||:||||||||| | |||||||
Db        322 TCACTACCTGTGGCTACA-GGTTTCCAAACCAAAGGCTCAGCTCTGCTCATACCAGGTTA 264

Qy       3127 AATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGC 3186
              ||||||||| :||||:||||| ||||||||||:|||||||||||||||:::||||| |
Db        263 AATACTTAGCGCTAAAATTATTCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAAC 204

Qy       3187 CTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAA 3246
              |||||||||||||||||||:||||||||||||||||||||: ::|||||||||:|||
Db        203 CTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAAGGTTCCTTGGCATAA 144

Qy       3247 YAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACA 3306
              :|||:|||||||||:||||||| | |||||:|| |:|||||||||| :|||||:|||||
Db        143 CAGGTTTCTGCCGAATATGGATT-CACAGGTATGACAAAATAGCCAGACCATTATATACA 85

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
              :||||||| ||||||||||||||||||||||||||:||||||||||:|:||||||:::||
Db         84 CTAATTAAAGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAA 25

Qy       3367 GTGGCTTTCCAGGCCCCTAA 3386
              |||||||||||| ||   ||
Db         24 GTGGCTTTCCAGACCNTAAA 5


RESULT 220
AI393478/c
LOCUS       AI393478                446 bp    mRNA    linear   EST 30-MAR-1999
DEFINITION  tg45g04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2111766 3', mRNA sequence.
ACCESSION   AI393478
VERSION     AI393478.1  GI:4223025
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1061   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 445.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2111766"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-PacI; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (fetal lung NbHL19W, testis NHT, and
                     B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made
                     in vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "

  Query Match             4.8%;  Score 356.2;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  385;  Conservative    9;  Mismatches   33;  Indels   10;  Gaps    1;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
              ||| ||| |||  ||||| ||| || ||||||||||||||||||||||||| ||||||:
Db        446 AATTTCAACTGTGCAACCCCTANTATGCCCCAATTCAGCAGGAAGCAGTTACAGCGGTCA 387

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
              || |||||||||||||||||||| |||||||||||||||  ||| ||||||||||
Db        386 TCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGAGGGCACTGAGAGACAA 327
```

```
Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             ||||||||||||||||||| ||:||||||||:||||| |||:||||||||||||  |
Db        326 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCGTAGCTGGGAAGGTGACTGC 267

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             |||||||||||||| |||||||| ||||||:|||||||  ||||||||||||||||||
Db        266 ATCCACCTTTAAACATGGGGCTTGCCACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 207

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||| |||||||||||||||||||||||||| :| ||||:||||
Db        206 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCACCTTTTGCCTGAG 147

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||||||||||||| |||||||||||||||  |  ||| ||||||||| ||||
Db        146 AGCACAGCAGGAGGGACAAGGATCGGGATATAAACCCGGGCATTCAAGCCGGCAATGGCA 87

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             |||||||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db         86 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGT----------TTTCACTCTAT 37

Qy       7566 TAAATCTTGCARCTGCR 7582
             ||||||||||:||||:
Db         36 TAAATCTTGCAACTGCA 20
```

RESULT 221
AQ774432
LOCUS AQ774432 476 bp DNA linear GSS 29-JUL-1999
DEFINITION HS_3198_B2_G04_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3198 Col=8 Row=N, genomic survey sequence.
ACCESSION AQ774432
VERSION AQ774432.1 GI:5654160
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 476)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3198 row: N column: 8
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 476.
FEATURES Location/Qualifiers
source 1. .476
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3198 Col=8 Row=N"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

Query Match 4.8%; Score 356.1; DB 1; Length 476;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 385; Conservative 15; Mismatches 52; Indels 3; Gaps 1;

```
Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
             || ||||: | :|||||||||||||:|||||| |||||||||| ||:|||||||:||||:|
Db          2 TATCTTCCAANGGAACACCTATCAAACATCATGAAGCCATTAAGAGATTATTATTGGCTG 61

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
             ||||||||||||:|||||||:|||||||||||||:||| |||||| ||||||||:|||
Db         62 TACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGATCATCAGAAAGGAAAGGAAA 121

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
             ||||||||:|| :|||:  |||||| :|||||||  | ||||||||||||||  |||||
Db         122 GGGAAATAGAATGGAACCACCAAGCGGATATTGAAACTAAAAGAGCTGCAAGGTGGGACC 181

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
             |||||||||||||||||  ||   ||| || ||| |||||||||| ||| ||||||||
Db         182 TTCCATTAGAAATGCTTATAGAAGGACCCCTACTATGGGGTAATCCCCTCCAGGAAACCA 241

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCC 4451
             |||||||||||||| ||| ||||||||||||||||||||||||||    ||||| ||||
Db         242 AGCCCCAGTACTCATCAGAAGAAACAGAATGGGGAACCTCACGAGGACAAAGTTTCCTCC 301

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
             ||||  || || || |||| |||||||| |||||||| ||||||||| ||| ||||||||
Db         302 CCTCACGATGGCTACCCACCGAAGAAGGAAAAATACTNTTGCCTGCAGCTAACCAATGGA 361

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
             ||||||||| ||||||||| || |||||||||||| |||| ||||||||||||  :||||
Db         362 AATTACTTANAACCCTTCACCAGACCTTTCACTTANGCATNGATAGCACCCATTGGATGG 421

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||||||| |||||||||||||||| || ||  ||
Db         422 CCAAATCATGATTTACTGGACCAGGACTNTTANAA 456
```

RESULT 222
BF992537
LOCUS BF992537 372 bp mRNA linear EST 23-JAN-2001
DEFINITION IL5-GN0178-301000-203-f06 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF992537
VERSION BF992537.1 GI:12398860
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 372)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-301000-203-f06&t3=2000-10-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 371.
FEATURES Location/Qualifiers

```
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.8%;  Score 355.8;  DB 1;  Length 372;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  351;  Conservative   14;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4479 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 4538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 60

Qy       4539 TTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCC 4598
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         61 TTCACTTAGGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCC 120

Qy       4599 TTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCC 4658
              ||||||||||||||||||:|||:|||||||||||||:|||||||:|:||||||||||||
Db        121 TTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCC 180

Qy       4659 TGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACT 4718
              |||||:|||||||||||||||||:||||||:||||||||||||||||:|:||:|||||
Db        181 TGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACT 240

Qy       4719 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 4778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGT 300

Qy       4779 ARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGT 4838
              |:||||||||||||||||||:|||||||| ||||||     |||||||||||||||||
Db        301 AGATACTTTCACGGGTTGGGCAGAGGCCTTCCNCTGTAGGGACAGAAAGGCCCAAGAGGT 360

Qy       4839 AATAAAGGCACT 4850
              |||||| |||||
Db        361 AATAAAAGCACT 372


RESULT 223
AI570707/c
LOCUS       AI570707                425 bp    mRNA    linear   EST 14-APR-1999
DEFINITION  tm79g09.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2164384 3'
            similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AI570707
VERSION     AI570707.1  GI:4534081
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1440   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 414.
FEATURES            Location/Qualifiers
     source         1. .425
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2164384"
                    /tissue_type="anaplastic oligodendroglioma"
                    /lab_host="DH10B"
                    /clone_lib="NCI_CGAP_Brn25"
                    /note="Organ: brain; Vector: pT7T3D-PacI; Site_1: Not I;
                    Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                    oligo(dT) primer [5'
                    TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                    T 3']; double-stranded cDNA was ligated to Eco RI
                    adaptors (Pharmacia), digested with Not I and cloned into
                    the Not I and Eco RI sites of  the modified pT7T3 vector.
                    Library is normalized, and was constructed  by Bento
                    Soares and M.Fatima Bonaldo."

  Query Match             4.7%;  Score 355.6;  DB 1;  Length 425;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  382;  Conservative    9;  Mismatches   28;  Indels   10;  Gaps    1;

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
              |||||||||| |||| | |||||| ||||||||||||||| |||||||||:||| | |||
Db        425 CTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCAGTCAA 366

Qy       7214 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 7273
              ||||| || |||||||| |||||||||||||||| |||| ||||||||||||||||||||
Db        365 CCTCCACAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCT 306

Qy       7274 GGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCT 7333
              |||||||||||||| ||:||||||||:||||||||:|||||||||||||| | |||||||
Db        305 GGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCT 246

Qy       7334 TTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 7393
              |||||||||||||||||||||||:||||||| ||||||||||||||||||||||||||||
Db        245 TTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGC 186

Qy       7394 TAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGC 7453
              |||||| ||||| |||||||||||||||||||||| ||||:||| ||:||||||||||||
Db        185 TAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGC 126

Qy       7454 AGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTT 7513
              | |||||||||||||||||||||||||||||| |  |||||||| |||||||| |||| |||
Db        125 AAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTT 66

Qy       7514 TGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTT 7573
              |||||||||||||||||||||||||| ||||          ||||||||||||||||||
Db         65 TGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTCACTCTATTAAATCTT 16

Qy       7574 GCARCTGCR 7582
              |||:||||:
Db         15 GCAACTGCA 7


RESULT 224
B17135/c
LOCUS       B17135                  491 bp    DNA     linear   GSS 04-JUN-1998
DEFINITION  344H17.TVB CIT978SKA1 Homo sapiens genomic clone A-344H17, genomic
            survey sequence.
ACCESSION   B17135
VERSION     B17135.1  GI:2124884
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: 344H17.TP 344H17.TPB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="A-344H17"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /clone_lib="CIT978SKA1"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A1"

  Query Match             4.7%;  Score 355.6;  DB 1;  Length 491;
  Best Local Similarity   79.6%;  Pred. No. 0;
  Matches  383;  Conservative   42;  Mismatches   52;  Indels    4;  Gaps    2;

Qy       2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             ||||||| |||||||:||:|::|||||| || |||||| ||| |  :|||  |||||||
Db        477 AGCCCAACACTTGAGCCAGTTCTCATACTTGAACACTCATGTTGTNTGGTACATGGATGA 418

Qy       3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
             ||||| | |:||:|||: | ||||| |||||||||||||||||||||||| :|| ||:||
Db        417 TTTACATCTAGCCGCCCCTGCAGAAGCCTTGTGCCATCAAGCCACCCAAGTGCTTTTAAA 358

Qy       3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
              |||||| |:|| |||||||||:|||||||||:|:|: ||||:||||||| | ||||
Db        357 CTTCCTCACCACTTGTGGCTACAAGGTTTCCAAACCAAAGTCTCAGCTCTGCTTATAGCA 298

Qy       3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
              |||||||||||||||: |||:|||||||||||||||||:||||||||| ||||||:::||
Db        297 GGTTAAATACTTAGGGTTAAAATTATCCAAAGGCACCAGGGCCCTCAGCGAGGAATGTAT 238

Qy       3182 CCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGG 3241
              ||| |||||||||||||||||| |||:||||||||| |||||| ||||:|: |||| |||
Db        237 CCATCCTATACTGGCTTATCCTGATCCCAAAACCCTGAAGCAATTAAGAGGTTTCCATGG 178

Qy       3242 CRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAW 3301
              |:|||:| |:||||||| ||:| ||||| |||||||:  ||:||| |||||||:|||||:
Db        177 CATAACAAGCTTCTGCCAAATACGGATT-CCCAGGTACAGCGAAAGAGCCAGGCCATTAT 119

Qy       3302 ATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAG 3361
              | |||:|||||||||||||||||||||||||||||| ||||||:|| |||||:|:||||||
Db        118 ACACACTAATTAAGGAAACTCAGAAAGCCAATACTCATTTAGTAGAATGGACACCTGAAG 59

Qy       3362 YMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTT 3421
              :::||   ||||| |||||||  ||  |    | | ||||||||:|||||||||||||:||
Db         58 CAGAAACAGCTTTTCAGGCCCTAAAGAA---GCCTTAACCCAAGCCCCAGTGTTAAGCTT 2

Qy       3422 G 3422
              |
Db          1 G 1
```

```
RESULT 225
AA552941/c
LOCUS       AA552941                431 bp    mRNA    linear   EST 09-SEP-1997
DEFINITION  nk61a10.s1 NCI_CGAP_Sch1 Homo sapiens cDNA clone IMAGE:1017978 3'
            similar to contains PTR5.b1 PTR5 repetitive element ;, mRNA
            sequence.
ACCESSION   AA552941
VERSION     AA552941.1  GI:2323195
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1994   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 390.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1017978"
                     /tissue_type="Schwannoma tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_Sch1"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Two
                     pooled bulk Schwannoma tumors. 5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size:  1.2 kb."

  Query Match             4.7%;  Score 355.3;  DB 1;  Length 431;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  383;  Conservative    7;  Mismatches   29;  Indels    9;  Gaps    1;

Qy       7131 CACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAG 7190
              |||||||| |||||||||| |||||| ||| |||||| |||||| |||| |||||||
Db        431 CACCCCTCCCAAGGAAATCTCAACTGCACCACCCCTACTATGCCCCAGTTCAACAGGAAG 372

Qy       7191 CAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGG 7250
              |||||||||||: |||||||||||||||| ||||||| |||||||||||||||| ||
Db        371 CAGTTAGAGCGGTCTTCGGCCAACCTCCCCAAAAGCACTTGGGTTTTCCTGTTGAGAGGG 312

Qy       7251 GGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGS 7310
              |||||||||||||||||||||||||||||||||||||||||||| ||:|||||||:|||||||:
Db        311 GGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGC 252

Qy       7311 TGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACC 7370
              |||||||||||| | |||| |||||||||||||||| ||||||| :|||||| |||||
Db        251 TGGGAAGGTGACCGCTTCCATCTTTAAACACGGGGCTTACAACTTAACTCACACATGACC 192

Qy       7371 AATCA---------GAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAA 7421
              |||||         |||||||||||||||||||||||||||  ||||||||||||||||
Db        191 AATCAGATAGTAAGGAGAGCTCACTAAAATGCTAATTAGGCAACAACAGGAGGTAAAGAA 132
```

```
Qy       7422 ATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACC 7481
              |||||||||||| :| ||||:|||||||||||| |||||||||||||||||||||||||||
Db        131 ATAGCCAATCACCTGTTGCCTGAGAGCACAGCGGGAGGGACAATGATCGGGATATAAACC 72

Qy       7482 CAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTC 7541
              || |  ||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db         71 CAGGCATTCAAGCCGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTC 12

Qy       7542 TGTTTTCA 7549
              ||||||||
Db         11 TGTTTTCA 4


RESULT 226
AA251070
LOCUS       AA251070                412 bp    mRNA    linear   EST 13-AUG-1997
DEFINITION  zs07d06.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684491 5'
            similar to SW:POL_FENV1 P31792 POL POLYPROTEIN ;, mRNA sequence.
ACCESSION   AA251070
VERSION     AA251070.1  GI:1886033
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1137   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 379.
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:5587845"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:684491"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from human tonsillar cells
                     enriched for germinal center B cells by flow sorting
                     (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
                     David Allman (NCI) and Dr. Gerald Marti (CBER).  cDNA
                     synthesis was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

  Query Match             4.7%;  Score 351.4;  DB 1;  Length 412;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  352;  Conservative   28;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
              |||| |||||| ||||||||||||||||||||||||||||||||||:||||:|||||||
Db          2 ATAAAGGAATAGTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 61

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
              ||||| ||:|||:| |||||||||||||||||||: |||||||:|||||||||||||||
Db         62 TAATAACCCTCACTCGGGCACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAG 121

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
              |||||:|:| ||||:|||||||| |||||||||||||::|||||||:|:|||||||||||
Db         122 ACTCTAAGTCTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAAT 181

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
              |||||| |||:|||:|||||||||||  :|||||||||||||||||||:|||| ||:|||
Db         182 TCCTAATTTCCGAGGGAACACCTATAGAACATCAGGAAGCCATTAGGAGATTACTATTGG 241

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
              |: |||||||||||:||||||||:|||||||||||| :||||||| || |||||||||:
Db         242 CTATACAGAAACCTAAAGAGGTGGCAGTCTTACACTGTTGGGGTCACCAGAAAGGAAAGG 301

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
              |||||||||||:|| :|||: ||||||| :|| || ||| :|||||||| ||||||  ||
Db         302 AAAGGGAAATAGAAAGGAACCGCCAAGCGGATCTTAAAGTCAAAAGAGCCGCAAGGATGG 361

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCT 4382
              |||||||||||||||||||||  ||   ||| ||||| ||||||||||||
Db         362 ACCCTCCATTAGAAATGCTTATGGAAGGGCCCCTAGTACAGGGTAATCCCT 412


RESULT 227
DQ033444
LOCUS       DQ033444                399 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Pan troglodytes HC3142 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ033444
VERSION     DQ033444.1  GI:66884653
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 399)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>399
                     /locus_tag="HC3142"

  Query Match             4.7%;  Score 350.7;  DB 1;  Length 399;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  360;  Conservative   20;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||| |||||||| |:|||||::|||||||||||||||||||||||
Db          1 AAGAGATCTTACTGTGTGGCCATCTCATAATGTGAACAGCATACTCACTGCTAAAGGAGA 60

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
              ||||||||||||||||||: ||| |||||||:|||||||||||||||||:|||||||
Db         61 CTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCAGT 120

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
```

```
                |||||:|||||||||||||||||||||||||||||:| ||||||||||||||||||||||
Db        121 GCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGA 180

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
                |||||||:||: ||||||||||||||:||||| |||||||||| ||  ||||||||||||
Db        181 AAAGATAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTGCTCGAGGGGACCT 240

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
                |:||||:|||||:||| ||||||||:|||| |||||||||||||||||||||| |||||||
Db        241 TCTAGAGGTTCCCTTGGCTGATCCCGACC-TCAACTTGTATACTGATGGAAGCTCCTTTG 299

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
                |||||||| |||||| ||| |:|| |||||||||||||||||| |||||||||:||||||
Db        300 TAGAAAAAAGACTTCAAAAGGCGGAGTATGCAGTGGTCAGTGACAATGGAATACTTGAAA 359

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGA 4029
                |||||||||||||||||||||||||| |||:|| |:|||||
Db        360 GTAATCCCCTCACTCCAGGAACTAGCGCTCAGGTGGCAGA 399


RESULT 228
AQ232635/c
LOCUS       AQ232635                456 bp    DNA     linear   GSS 26-SEP-1998
DEFINITION  HS_2026_A1_A08_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2026 Col=15 Row=A, genomic survey
            sequence.
ACCESSION   AQ232635
VERSION     AQ232635.1  GI:3657864
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2026  row: A  column: 15
            Class: BAC ends
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2026 Col=15 Row=A"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.7%;  Score 350.7;  DB 1;  Length 456;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  378;  Conservative   42;  Mismatches   36;  Indels    3;  Gaps    3;

Qy       2957 YCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGC 3016
                :||:|::| |||||||||| ||||||| ||||:|||  ||| ||||| | ||||:||:
Db        456 CCAATTCTTATACCTGGACCCTCTTGTTCTTCAGTACATGGGTGATTGAGTTTTAGCCAA 397

Qy       3017 CYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTG 3076
                |::|||||||||  |||||||| |||||| |||||:|||||:|||||| |  |:|||||
Db        396 CCATTCAGAAACTNTGTGCCATCGAGCCACTCAAGCGCTCTTAAATTTCATGCCTACCTG 337

Qy       3077 TGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGG 3136
                ||||||||:|||| ||||||:|:|:|||| :|||||| ||||||||||||||||||||
Db        336 TGGCTACAAGGTTGCCAAACCAAAGGCTCGGCTCTGCTTACAGCAGGTTAAATACTTAGG 277

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
                : || :|||||||||||||||:|||||||||||||||:::|||||||||||| ||||
Db        276 GTTAGAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATAGTGGC 217

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTG 3256
                ||||||||||:| |||||||||||||||||:|::|||||||||:|| :|||:|||||
Db        216 TTATCCTCATCCCCAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAGCAGGTTTCTG 157

Qy       3257 CCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGG 3316
                | |||:|||||| |||||||: |||:|||||||||| :|||||:|||| |||||||||
Db        156 CTGAATATGGATT-CCCAGGTACGGCAAAATAGCCAGACCATTATATACATTAATTAAGG 98

Qy       3317 AAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCC 3376
                |||||||||||||| ||||||||||:|||||||||:|: |||||:::||| ||| ||||
Db         97 AAACTCAGAAAGCCGATACCCATTTAGTAAGATGGATACGTGAAGCAGAAGAGGCATTCC 38

Qy       3377 AGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTT 3415
                ||| || |||||||||||||| ||| ||||: ||||||||
Db         37 AGG-CCATAAAGAAGGCCTT-AACACAAGCTCCAGTGTT 1


RESULT 229
AG196511
LOCUS       AG196511                440 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-076D12.T7, genomic survey
            sequence.
ACCESSION   AG196511
VERSION     AG196511.1  GI:45228687
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 440)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-076D12.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"
```

```
  Query Match             4.7%;  Score 350.1;  DB 1;  Length 440;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  376;  Conservative   14;  Mismatches   39;  Indels    3;  Gaps    1;

Qy       4206 TAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAG 4265
             ||:||||:||| |||||||||:| |||||: |||||||||||||:|||||||| ||||
Db             9 TACTGGCTGTATAGAAACCTAAATAGGTGGCTGTCTTACACTGCTGGGGTCATCAGAAAG 68

Qy       4266 GAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAA 4325
             |||||:||||| |||||:|||:|||:|||||||| :|||||||||:|||||||| ||||
Db            69 GAAAGGAAAGGAAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAA 128

Qy       4326 GGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCC 4385
             |||||||| |||||||||||||||||||  ||   ||| |||||| |||||||||| |||
Db           129 GGCAGGACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCC 188

Qy       4386 GGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CA 4442
             |||||||||||||||| ||||||||||||||  |||||||||||||||||| |||     |
Db           189 GGGAAACCAAGCCCCAATACTCAGCAGGAGAAGTAGAATGGGGAACCTCACAAGGACATA 248

Qy       4443 GTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTA 4502
             |||||||||||||| ||| || |||||||||||||||  ||||||| ||||||||| |||
Db           249 GTTTTCTCCCCTCAGGATGGCTAGCCACTGAAGAAGGAGAAATACTGTTGCCTGCAGCTA 308

Qy       4503 TCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCC 4562
              |||||||||||||||||||||||||||| ||||||||||||||||||| ||||||||
Db           309 ACCAATGGAAATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACTG 368

Qy       4563 ATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAK 4622
             ||||:|||||||||| |||||||||||||  |||||||||||||||||||| ||:|||:
Db           369 ATCAGATGGCCAAATTATTATTTACTGGACTGGGCCTTTTCAAAACTATCAAACAGATAG 428

Qy       4623 TCAGGGCCTGTG 4634
             ||||||||  ||
Db           429 TCAGGGCCCATG 440


RESULT 230
BF993782
LOCUS       BF993782                366 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-011100-205-d06 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF993782
VERSION     BF993782.1  GI:12400105
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            011100-205-d06&t3=2000-11-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 366.
FEATURES            Location/Qualifiers
     source         1. .366
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="GN0178"
                    /note="Organ: placenta_normal; Vector: puc18; Site_1:
                    SmaI; Site_2: SmaI; A mini-library was made by cloning
                    products derived from  ORESTES PCR (U.S. Letters Patent
                    application No. 196,716 -  Ludwig Institute for Cancer
                    Research) profiles into  the pUC 18 vector. Reverse
                    transcription of tissue mRNA and cDNA amplification were
                    performed under low stringency conditions."

  Query Match             4.7%;  Score 350;  DB 1;  Length 366;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  350;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4839 AATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAG 4898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db              5 AATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAG 64

Qy       4899 TGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTAT 4958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             65 TGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTAT 124

Qy       4959 ACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAA 5018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            125 ACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAA 184

Qy       5019 TGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGY 5078
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db            185 TGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGC 244

Qy       5079 TCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCC 5138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            245 TCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCC 304

Qy       5139 ATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCC 5192
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db            305 ATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTCCCCC 358


RESULT 231
H01372
LOCUS       H01372                  472 bp    mRNA    linear   EST 19-JUN-1995
DEFINITION  yi99e01.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:147384 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
            POLYPROTEIN - ;contains PTR7 repetitive element ;, mRNA sequence.
ACCESSION   H01372
VERSION     H01372.1  GI:864305
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
```

```
            Email: est@watson.wustl.edu
            Insert Size: 790
            High qality sequence stops: 384
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 790   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 384.
FEATURES              Location/Qualifiers
     source           1. .472
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:559031"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:147384"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares placenta Nb2HP"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.7%;  Score 349.3;  DB 1;  Length 472;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  379;  Conservative    7;  Mismatches    5;  Indels    3;  Gaps    3;

Qy       6863 KAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGA 6922
              :| |   :||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db          7 GAACGAGGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGA 66

Qy       6923 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA 6982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA 126

Qy       6983 TCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||:|||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        127 TCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGG 186

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCT-GGACCGGCCTGYTAGC 7101
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||:||||
Db        187 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGGACCGGCCTGCTAGC 246

Qy       7102 CCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAA 7161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAA 306

Qy       7162 CCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTA-GAGCGGTSGTCGGCCAACCTCCCC 7220
              |||||||||||||||||||||||||||||||||| |||||||:|||||||||||| |||
Db        307 CCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGGAGCGGTCGTCGGCCAACCTNCCC 366

Qy       7221 AACAGCACTTA-GGTTTTCCTGTTGAGATGGGGG 7253
              |||||||||||| |||||||||||||||||||||
Db        367 AACAGCACTTAGGGTTTTCCTGTTGAGATGGGGG 400


RESULT 232
AQ609759/c
LOCUS       AQ609759                481 bp    DNA     linear   GSS 15-JUN-1999
DEFINITION  HS_5076_A1_C09_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=652 Col=17 Row=E, genomic survey sequence.
ACCESSION   AQ609759
VERSION     AQ609759.1  GI:5071035
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 652  row: E  column: 17
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 481.
FEATURES              Location/Qualifiers
     source           1. .481
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=652 Col=17 Row=E"
                      /sex="male"
                      /clone_lib="RPCI-11 Human Male BAC Library"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      Male blood DNA was isolated from one randomly chosen donor
                      and partially digested with a combination of EcoRI and
                      EcoRI Methylase. Size selected DNA was cloned into the
                      pBACe3.6 vector at EcoRI sites"

  Query Match             4.7%;  Score 348.7;  DB 1;  Length 481;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  367;  Conservative   24;  Mismatches   43;  Indels    1;  Gaps    1;

Qy       3703 GACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGC 3762
              ||||||: ||| ||||||    ||||| ||||| |||||:| | ||||||| :| || ||
Db        481 GACAACCATTTGCTTAAAATCAAGGCTTTATTAATTGAAGGACTAGTGCTGAGATTGCGC 422

Qy       3763 ACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYAT 3822
              ||||||||||||||||||||:| ||||||||||||||||||||||||||| |:||: |
Db        421 ACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGAGAGAACGT 362

Qy       3823 AACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCY 3882
              |||||||||| :| ||| |||||||||||| ||||| | |||||||||:||||: ||||:
Db        361 AACTGTCAACAGGTGATTGCTCAAACCTATGTCACTCAAAGGGACCTTCTAGAGATTCCC 302

Qy       3883 TTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACT 3942
              ||||||||||: ||| ||||||||||||||||| |||||||||||||| |||||||||
Db        301 TTGACTGATCCCTACC-TCAACTTGTATACTGATGAAAGTTCCTTTGTAGCAAAAGGACT 243

Qy       3943 TCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCAC 4002
              | |||||:|||||||||||||||||||||||||||||:||||||||||||||||||||
Db        242 TTGAAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCAC 183

Qy       4003 TCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGG 4062
              | |||||||||| |||:|| |:|||||||||||||:|||:| |||||||||||||| |
Db        182 TGCAGGAACTAGCGCTCAGTTGGCAGAACTAATAGCCCTCATTCGGGCACTAGAATTAAG 123

Qy       4063 AGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGC 4122
              | |||:  ||||||:|||||||||||||| |||:|:||||||:|| ||| | ||| ||
Db        122 AAAAGGNNAAAGGGTAAATATATATACAGAATCTAAGTATGCTTACNTAGCCCTCCNAGC 63
```

```
Qy       4123 CCATGMRGCAATATG 4137
                ||| ::||||||||
Db         62 CCACACAGCAATATG 48


RESULT 233
AQ132081/c
LOCUS       AQ132081                428 bp    DNA     linear   GSS 23-SEP-1998
DEFINITION  HS_2166_B2_D12_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2166 Col=24 Row=H, genomic survey
            sequence.
ACCESSION   AQ132081
VERSION     AQ132081.1  GI:3509247
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 428)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2166  row: H  column: 24
            Class: BAC ends
            High quality sequence stop: 428.
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2166 Col=24 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.7%;  Score 348.6;  DB 1;  Length 428;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  368;  Conservative   39;  Mismatches   21;  Indels    4;  Gaps    3;

Qy       2984 CCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGC 3043
                ||||:||| | ||||||||||||||:||:|||::|||||||||||||||| |||||
Db        428 CCTTCGGTACGCGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCACCAAGC 369

Qy       3044 CACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGC 3103
                |||||||||:|||||:||||||||||:|||||||||||||:||||||||||:|:|:||
Db        368 CACCCAAGCGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAACCAAAGGC 309

Qy       3104 TCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGC 3163
                |||:|||||||||||||| ||||||||||||:||||:|||||||||||||||||:|||
Db        308 TCAGCTCTGCTCACAGCAGTTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGC 249

Qy       3164 CCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCA 3223
                |||||||| ||||:::|||||||||||||| ||||||||||||:| ||||||||||||
Db        248 CCTCAGTGGGGAATGTATCCAGCCTATACT-GCTTATCCTCATCCCCAAACCCTAAAGCA 190

Qy       3224 ACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCR 3283
                ||||||:|::|||||||||:|||:|||:||||||||||:||||||| | |||||: |||:
Db        189 ACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-CNCAGGTACGGCA 131

Qy       3284 AAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAR 3343
                |||||||||| :|||||:|||||:|||||||||||||||||||||||||||||||||:
Db        130 AAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAG 71

Qy       3344 TAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCA 3403
                |||||||||:|:||||||:::||||||||||  |||||   ||  | | | |  ||| ||
Db         70 TAAGATGGACACCTGAAGCAGAAGTGGCTTTAAAGGCCATAAAGAAGGCCAT--AACNCA 13

Qy       3404 AGYCCCAGTGTT 3415
                ||:| |||||||
Db         12 AGCCNCAGTGTT 1


RESULT 234
R78828
LOCUS       R78828                  464 bp    mRNA    linear   EST 09-JUN-1995
DEFINITION  yi90c04.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:146502 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
            POLYPROTEIN - ;contains PTR7 repetitive element ;, mRNA sequence.
ACCESSION   R78828
VERSION     R78828.1  GI:855109
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 793
            High quality sequence stops: 353
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 793   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 353.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:558051"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:146502"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.7%;  Score 348.6;  DB 1;  Length 464;
  Best Local Similarity   94.7%;  Pred. No. 0;
```

```
  Matches  372;  Conservative    7;  Mismatches   12;  Indels    2;  Gaps    2;

Qy       6863 KAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGA 6922
              :| |   :|||||||||||||||||||||||||||||||||| ||||:|||||||||||||||
Db          8 GAACGAGGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTTAGCCAATGGATGCCCTGGA 67

Qy       6923 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA 6982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA 127

Qy       6983 TCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||:|||||||||||||||||||||||| ||||||||||||||:|||||||||||||
Db        128 TCTTTAACCTCCTTGTTAACTTTGTCTCTTNCAGAATCGAAGCTGTAAAACTACAAATGG 187

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        188 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCC 247

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 307

Qy       7163 CTCTAC-TACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG-TCGGCCAACCTCCCC 7220
              |||||| ||||||||||||||||||||||||||||||||||:| ||||||||||||||||
Db        308 CTCTACTTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTTCGGCCAACCTCCCC 367

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGG 7253
              |||||||||||||||||||||||      ||||
Db        368 AACAGCACTTAGGTTTTCCTGTTTGAGATGGGG 400


RESULT 235
AV708482
LOCUS       AV708482                 432 bp    mRNA    linear   EST 09-OCT-2000
DEFINITION  AV708482 ADC Homo sapiens cDNA clone ADCAOB08 5', mRNA sequence.
ACCESSION   AV708482
VERSION     AV708482.1  GI:10725747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 432)
  AUTHORS   Peng,Y., Song,H., Huang,Q., Huang,C., Gu,Y., Yang,Y., Gao,G.,
            Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z.,
            Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S.,
            Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA ADC clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .432
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADCAOB08"
                     /tissue_type="Adrenal gland"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="ADC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"

  Query Match             4.6%;  Score 345.5;  DB 1;  Length 432;
  Best Local Similarity   88.5%;  Pred. No. 0;

  Matches  377;  Conservative    5;  Mismatches   43;  Indels    1;  Gaps    1;

Qy         93 TTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGC 152
              |||||||||||||| ||||||  ||||| |||||||||||| ||| ||||||||| |||
Db          3 TTTCACTCTATTAAGTCTTGCTTCTGCAGTCTTCTGGTCCATATTTGTTACGGCTCCAGC 62

Qy        153 TGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACT 212
              |||||||||||||| |||||||||  ||| ||||||||  |    |||| ||| ||||||||||
Db         63 TGAGCTTTTGCTCGCCGTCCACCCCTGTTGTTTGCCGCTGTCGCAGACCCGCCGCTGACT 122

Qy        213 CCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATT 272
              |||||||||  |||||  |||||||||| |||||||||||||||||||:|||||||||
Db        123 TCCATCCCTCCGGATCCAGCAGGGTGTCTGCTGTGCTCCTGATCCAGCGAGGCGCCCATT 182

Qy        273 GCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTT 332
              |||  ||||  || ||  ||  |||||||||||||| |||||||||||||| |||||||||
Db        183 GCCACTCCTGATCGGACTGAAGGCTTGCCATTGTTCCTGCACGGCTAAGAGCCTGGGTTC 242

Qy        333 GTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGC 392
              ||  |||||  |||||||||||| |||||||||||| ||||||||| ||||||||||||
Db        243 GTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGC 302

Qy        393 TTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGA 452
              ||||||||:|  |||||||||| ||| ||||||||||||||||||||||||||||||||
Db        303 TTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATCCGTGA 362

Qy        453 GGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCT 512
              ||:||| |||| |||||||||||| ||||:|||||||||| |||||||||| ||||||
Db        363 GGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCGTCTTGGAAGCAGCCTGCC 421

Qy        513 ACCRTC 518
              || :||
Db        422 ACTATC 427


RESULT 236
D78692
LOCUS       D78692                   379 bp    mRNA    linear   EST 09-FEB-1996
DEFINITION  HUM503H08B Human placenta polyA+ (TFujiwara) Homo sapiens cDNA
            clone GEN-503H08 5', mRNA sequence.
ACCESSION   D78692
VERSION     D78692.1  GI:1180565
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
            Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
            Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
            Maekawa,H., Shin,S. and Nakamura,Y.
  TITLE     Fujiwara et al. (1995)
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Tsutomu Fujiwara
            Otsuka GEN Research Institute
            Otsuka Pharmaceutical Co.,Ltd
            463-10 kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
            Tel: 0886-65-2888
            Fax: 0886-37-1035.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GEN-503H08"
                     /tissue_type="placenta"
                     /clone_lib="Human placenta polyA+ (TFujiwara)"

  Query Match             4.6%;  Score 345;  DB 1;  Length 379;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  369;  Conservative    2;  Mismatches    7;  Indels    2;  Gaps    2;
```

```
Qy       5688 GTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTS 5747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db          1 GTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTC 60

Qy       5748 TAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 5807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 120

Qy       5808 CACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTG 5867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTG 180

Qy       5868 TCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGA 5927
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        181 TCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGG-ATGTCTGA 239

Qy       5928 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACT 5987
             ||||||||||||||||||||||||||||||||||||    |||  ||||||||||||||||
Db        240 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATTGTAAAAGAGTAATCTCCCAACT 299

Qy       5988 CA-CCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATG 6046
              || ||:||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        300 CACCCGGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATG 359

Qy       6047 AAACCCTCCGTACCCATACT 6066
              |||||||||||||| |||||
Db        360 AAACCCTCCGTACCGATACT 379
```

```
RESULT 237
AI379210/c
LOCUS       AI379210                433 bp    mRNA    linear   EST 28-MAR-1999
DEFINITION  td01g11.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2074436 3',
            mRNA sequence.
ACCESSION   AI379210
VERSION     AI379210.1  GI:4189063
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
            M.D., Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1011   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2074436"
                     /tissue_type="B-cell, chronic lymphotic leukemia"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_CLL1"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTGCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
```

```
  Query Match             4.6%;  Score 344.4;  DB 1;  Length 433;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  369;  Conservative    9;  Mismatches   25;  Indels   10;  Gaps    1;

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             | |||||| |||||||||||||| |||||||||:||| | |||||||||||| |||||||
Db        432 ATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCAGTCAACCTCCCCAGCAGCACT 373

Qy       7230 TAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGA 7289
             | |||||||||||||||| |||| ||||||||||||||||||||||||||||||||| ||
Db        372 TGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGA 313

Qy       7290 YTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTG 7349
             :|||||||||:|||||||||:|||||||||||||| | ||||||||||||||||||||||
Db        312 CTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTG 253

Qy       7350 CAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACA 7409
             ||||||||:||||||| |||||||||||||||||||||||||||||||| ||||| |||
Db        252 CAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTAATTAAGCAAAAACA 193

Qy       7410 GGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATC 7469
             |||||||||||||||||| ||||:||| ||:|||||||||||| ||||||||||||||||
Db        192 GGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGCAAGAGGGACAATGATC 133

Qy       7470 GGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTT 7529
             ||||||||||||||  |  |||||||| |||||||| |||| ||||||||||||||||||
Db        132 GGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTT 73

Qy       7530 GTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||| ||||          |||||||||||||||||||:||||:
Db         72 GTATGGGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 30
```

```
RESULT 238
N41878
LOCUS       N41878                  468 bp    mRNA    linear   EST 24-JAN-1996
DEFINITION  yw72g06.r1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:257818 5' similar to contains PTR7.t1 PTR7 repetitive
            element ;, mRNA sequence.
ACCESSION   N41878
VERSION     N41878.1  GI:1165909
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 308
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
```

```
          High quality sequence stop: 308.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3887428"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:257818"
                     /dev_stage="two placentae: one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."

  Query Match             4.6%;  Score 341.8;  DB 1;  Length 468;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  390;  Conservative    7;  Mismatches   17;  Indels    8;  Gaps    4;

Qy       6927 CCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTT 6986
             ||||||||  ||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 CCCCTTCTTNNGACCTCTAGCAGCTANAATATTGCTACTCCTCTTTGGACCCTGTATCTT 60

Qy       6987 TRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCC 7046
             |:|||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db         61 TAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCC 120

Qy       7047 CAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACG 7106
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        121 CAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACG 180

Qy       7107 ATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCT 7166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCT 240

Qy       7167 ACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGC 7226
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        241 ACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGC 300

Qy       7227 ACTTAGGTTTTCCTG-TTGAGATGGGGGACTGA-GAGACAGGACTAGCTGGATTTCC-TA 7283
             |||||||||||||||| |||||||||||||||| |||||||||||||||||||||||| ||
Db        301 ACTTAGGTTTTCCTGTTTGAGATGGGGGACTGAGGAGACAGGACTAGCTGGATTTCCTTA 360

Qy       7284 GGCTGAYTAAGAATCCYTAAGCC-----TAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||:|||||||||:|||  |       |: |    |  |||||||||||||||  |
Db        361 GGCTGACTAAGAATCCCTAAAGCCTTAGCTGGGGAAGGGTGGACCACATCCACCTTTTTA 420

Qy       7339 CA 7340
                ||
Db        421 AA 422


RESULT 239
R27412/c
LOCUS       R27412                   424 bp    mRNA    linear   EST 24-APR-1995
DEFINITION  yh46d11.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:132789 3', mRNA sequence.
ACCESSION   R27412
VERSION     R27412.1  GI:783547
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 695
            High quality sequence stops: 367
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 695   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 367.
FEATURES             Location/Qualifiers
     source          1. .424
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:538431"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:132789"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.6%;  Score 341.6;  DB 1;  Length 424;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  385;  Conservative    8;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       7158 ACAACCTCTACTACGCCCCAATT-CAGCAGGAAGCAGTTAGAG-CGGTSGTCGGCCAACC 7215
             | | || ||||||  ||||| || |||||||||||||||||| | ||:||||  |||||
Db        423 AAACCCCCTACTAGCCCCCAGTTCCAGCAGGAAGCAGTTAGAGCCAGTCGTCGATCAACC 364

Qy       7216 TCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGG 7275
             |||||||||||||||| |||||||||  |||||| ||||||||||||||||||||||||||
Db        363 TCCCCAACAGCACTTGGGTTTTCCTATTGAGAGGGGGGACTGAGAGACAGGACTAGCTGG 304

Qy       7276 ATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTT 7335
             ||| |||||||  ||:|||||||||:|||||||||:|||||||||||  ||||||||||
Db        303 ATTCCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTT 244

Qy       7336 AAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTA 7395
             ||||||||||||||||||||||:|||||||  |||||||  ||||||||||||||||||||
Db        243 AAACACGGGGCTTGCAACTTAGCTCACACCCAACCAATC--AGAGCTCACTAAAATGCTA 186

Qy       7396 ATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAG 7455
             |||||||||| ||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db        185 ATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACAGCAG 126

Qy       7456 GAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTG 7515
             ||||| ||| |||||||||||||||||| |  |||||||||||||||||||| ||||||||
Db        125 GAGGGGCAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGGCAAACCCCTTTG 66

Qy       7516 GGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGC 7575
             |||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db          65 GGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGC 6
```

```
Qy       7576 ARCT 7579
              |:||
Db          5 AACT 2
```

RESULT 240
AQ210561
LOCUS AQ210561 408 bp DNA linear GSS 18-SEP-1998
DEFINITION HS_2212_A2_E02_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2212 Col=4 Row=I, genomic survey sequence.
ACCESSION AQ210561
VERSION AQ210561.1 GI:3619530
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2212 row: I column: 4
Class: BAC ends
High quality sequence stop: 408.
FEATURES Location/Qualifiers
source 1. .408
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2212 Col=4 Row=I"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

Query Match 4.6%; Score 340.8; DB 1; Length 408;
Best Local Similarity 84.8%; Pred. No. 0;
Matches 347; Conservative 33; Mismatches 27; Indels 2; Gaps 1;

```
Qy       3422 GCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAG 3481
              | |  |:| || ||||||| :||:|||::|||||  || |||||:|||:|||||:||||
Db          1 GGCCTCAGTGCTAGACTTTNCTTTATATGTCACAG--AATAACAGGAATAGCTCTAGGAG 58

Qy       3482 TCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTG 3541
              ||||||||||:|| :|||||: ||||||||||::||||::|| |||:||||||||:||
Db         59 TCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAAGGAAACTG 118

Qy       3542 ATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAG 3601
              |||||||| |||||||||:|:|||||||||: |||||||  ||||||||||::|||
Db        119 ATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGTGGCAGCAGTAGCAGTCTTAG 178

Qy       3602 TATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKG 3661
              |||||||||||||||||| ||||||||:|||||||||||||||||||||||||||||:|
Db        179 TATCTGAAGCAGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATG 238

Qy       3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
              ||||::|||||||||||||||||||||||||||||||||||||||: ||| |||||||
Db        239 TGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTGCTTAAAT 298

```
Qy       3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
              :|||||||||||||| |||||:| |||||||||:|||| |||||||||||||||||| |
Db        299 ATCAGGCTCTATTAGTTGAAGGACCAGTGCTGCGACTGCACACTTGTGCAACTCTTAATC 358

Qy       3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCA 3830
              |||:| ||||| |||||| |||||||||||| | |:||:||||||||||
Db        359 CAGCCACATTTTTTCCAGGCAATGAAGAAAAAACAGAACATAACTGTCA 407
```

RESULT 241
AQ734369
LOCUS AQ734369 402 bp DNA linear GSS 15-JUL-1999
DEFINITION HS_3112_B1_G02_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3112 Col=3 Row=N, genomic survey sequence.
ACCESSION AQ734369
VERSION AQ734369.1 GI:5505921
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3112 row: N column: 3
Seq primer: T7
Class: BAC ends
High quality sequence stop: 402.
FEATURES Location/Qualifiers
source 1. .402
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3112 Col=3 Row=N"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

Query Match 4.5%; Score 339.7; DB 1; Length 402;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 349; Conservative 27; Mismatches 26; Indels 1; Gaps 1;

```
Qy       3463 ACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCR 3522
              ||||:||:||||||:||||||||||||||:|||:|||||:|||||||||||||::||||:
Db          1 ACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCCGTGGCA 60

Qy       3523 YACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTG 3582
              :||||||:||||||||:|||||||||||||| ||||| |:||||||||||:||||||||
Db         61 TACCTGAGTAAGGAAATTGATGTAGTGGCAAA-GGTTGTCCTCATTGTTTATGGGTAGTG 119

Qy       3583 GTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACT 3642
              ||||| |||||::|||||||||||||||||||||| ||||||||:||||||||||||
Db        120 ACGGCAGCAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGATCTTACT 179

Qy       3643 GTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCA 3702
```
```

```
                ||||||||||||||||||: ||||::|||||||||||||||||| |||||||||||||||||
Db        180 GTGTGGACATCTCATGATATGAACGGCATACTCACTGCTAAAAGAGACTTGTGGCTGTCA 239

Qy       3703 GACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGC 3762
                ||||||  ||||||||||:||| |||||  ||||||||:|||||||||||:|||||||
Db        240 GACAACGATTTACTTAAATATCAAGCTCTTGTACTTGAAGGGCCAGTGCTGCGACTGTGC 299

Qy       3763 ACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYAT 3822
                ||||||||||||||||||  ||:| ||||  ||||| |||||||||||||||| :||:||
Db        300 ACTTGTGCAACTCTTAACTGAGCCACATTGTTTCCACACAATGAAGAAAAGATTGAACAT 359

Qy       3823 AACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
                ||||||||||:||||| |||| |||| |||   ||||||||
Db        360 TACTGTCAACAAGTAATTGCTCAGACCTGTGCTGTTCGAGGGG 402


RESULT 242
BI049622
LOCUS       BI049622                362 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  CM3-GN0297-020101-591-d08 GN0297 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI049622
VERSION     BI049622.1  GI:14457152
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 362)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-GN0297-
            020101-591-d08&t3=2001-01-02&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 3
            High quality sequence stop: 362.
FEATURES             Location/Qualifiers
     source          1. .362
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0297"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.5%;  Score 339.5;  DB 1;  Length 362;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  350;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       5063 ACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAA 5122
                ||| | ||  | |||:|||| ||||||||||||||||||||||||||||||||||||
Db          3 ACCACTCAATGGCTGCTCTG-TGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAA 61

Qy       5123 AAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTG 5182
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTG 121

Qy       5183 TGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACA 5242
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACA 181

Qy       5243 AGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCC 5302
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCC 241

Qy       5303 WWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTG 5362
                ::|||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        242 TTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTG 301

Qy       5363 GGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGA 5422
                ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        302 GGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGA 361

Qy       5423 G 5423
                |
Db        362 G 362


RESULT 243
R76136
LOCUS       R76136                  378 bp    mRNA    linear   EST 06-JUN-1995
DEFINITION  yi71b03.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:144653 5' similar to contains PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   R76136
VERSION     R76136.1  GI:850818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 597
            High quality sequence stops: 279
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 597   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 279.
FEATURES             Location/Qualifiers
     source          1. .378
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:562655"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:144653"
                     /sex="Female"
```

/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta Nb2HP"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. "

Query Match 4.5%; Score 339.5; DB 1; Length 378;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 358; Conservative 7; Mismatches 5; Indels 3; Gaps 2;

```
Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db          1 ATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAA 60

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 120

Qy       7189 AGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAT 7248
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAT 180

Qy       7249 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTA 7308
             ||||||||||||||||||||||||||||||||||||||||:||||||||:||||||||
Db         181 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTA 240

Qy       7309 GSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG--CTTGCAACTTAGYTCACACCT 7366
             |:|||||||||||||||||||||||||||||||||||  ||||||||||:||||||||
Db         241 GCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGGCTTTGCAACTTAGCTCACACCT 300

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAA-GAAATAG 7425
             ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||| |||
Db         301 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGTAAAGGAANTAG 360

Qy       7426 CCAATCATYTATT 7438
             ||| ||||:| ||
Db         361 CCANTCATCTTTT 373
```

RESULT 244
N55091/c
LOCUS N55091 385 bp mRNA linear EST 28-JAN-1997
DEFINITION yv43e03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:245500 3', mRNA sequence.
ACCESSION N55091
VERSION N55091.1 GI:1197970
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 385)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 907 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 276.
FEATURES Location/Qualifiers
source 1. .385
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3794746"
/db_xref="taxon:9606"
/clone="IMAGE:245500"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

Query Match 4.5%; Score 339; DB 1; Length 385;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 360; Conservative 9; Mismatches 10; Indels 2; Gaps 2;

```
Qy       7202 GTSGTCGGCCAACCTCCCCCAACAGCACTTAGGTTTTCCTGTTG-AGATGGGGGACTGAGA 7260
             ||:| |    | |  ||||||||||||||||||||||||||| ||||||||||||||||
Db         380 GTCGGCCCAAACCTCCCCCAACAGCACTTAGGTTTTCCTGTTGAAGATGGGGGACTGAGA 321

Qy       7261 GACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
             ||||||||||||||||||||||||||||:||||||||:|| ||||||:||||||||||
Db         320 GACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAGGCCTAGCTGGGAAGGTG 261

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db         260 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAG 201

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
             |||||||||||||||||||||||||| |||||||||||||||||||:||||||
Db         200 CTCACTAAAATGCTAATTAGGCAAAGACANGAGGTAAAGAAATAGCCAATCATCTATTGC 141

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
             :|||||||||||||||||||||||||||||||||||||||||||||:||||| |||||||
Db         140 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGA-CCGGCAA 82

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 22

Qy       7561 TCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||:||||
Db          21 TCTATTAAATCTTGCAACTGC 1
```

RESULT 245
AA250958/c
LOCUS AA250958 410 bp mRNA linear EST 13-AUG-1997
DEFINITION zs07d10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684499 3' similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA sequence.
ACCESSION AA250958
VERSION AA250958.1 GI:1885919
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1390   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .410
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:684499"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from human tonsillar cells
                     enriched for germinal center B cells by flow sorting
                     (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
                     David Allman (NCI) and Dr. Gerald Marti (CBER).  cDNA
                     synthesis was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             |||||||||||||||  |||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:|||||||||||||| |||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  ||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             |||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||||  |||||||||||||||||||||||:|||
Db         50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 246
BX388918/c
LOCUS       BX388918                 403 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX388918 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 3-PRIME, mRNA sequence.
ACCESSION   BX388918
VERSION     BX388918.2  GI:46845082
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30461046.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAB005ZC06_CS00428_2&c=4215.r
            .
FEATURES             Location/Qualifiers
     source          1. .403
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YJ18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             4.5%;  Score 338.8;  DB 1;  Length 403;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  361;  Conservative    8;  Mismatches   10;  Indels    2;  Gaps    2;

Qy       6562 GGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGT-ATCACAACCTCTACTCAGTTCTA 6620
             |||||||||||||| |||||  | ||||||||||||  ||| |||||| ||||||||| |
Db        398 GGAGTGCTAGGTGCGCTAGGCGCAGGCATTGGCGGTGGTCAGAACCTGTACTCAGTTATG 339

Qy       6621 CTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCAC 6680
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        338 CTACAAACTATCTCGAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCAC 279

Qy       6681 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGA 6740
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        278 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGA 219

Qy       6741 CTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 6800
             ||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        218 CTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTA 159

Qy       6801 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACG 6860
             |||||||||||||||||||||||||||||||||||:|||||||||:|||||||||||||||
Db        158 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACG 99

Qy       6861 TAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTG 6920
             ||:||||||:||||||||||||||||||||||||||||||||||||:|||||||||||| |
Db         98 TAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCC-G 40

Qy       6921 GATTCTCCCCTTCTTAGGACC 6941
             |||||||||||||||||||||
Db         39 GATTCTCCCCTTCTTAGGACC 19
```

```
RESULT 247
BI056592/c
LOCUS       BI056592                345 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-210201-013-E03 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056592
VERSION     BI056592.1  GI:14464122
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            210201-013-E03&t3=2001-02-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 344.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.5%;  Score 338.4;  DB 1;  Length 345;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  339;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6029 ATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATA 6088
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        345 ATCTCTCAAAACTACATGAAACCCTCCGTAACCATACTCGCCTGGTAAGCCTATTTAATA 286

Qy       6089 CCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 6148
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        285 CCACCCTCACTGGGCTCCATAAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCC 226

Qy       6149 TCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 6208
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        225 TCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCA 166

Qy       6209 GCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAA 6268
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        165 GCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAA 106

Qy       6269 CCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCC 6328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 ACCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCC 46

Qy       6329 AATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTA 6372
              |||||||||||||||||||||||||||||||||||||||||||
Db          45 AATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTA 2


RESULT 248
AI288235/c
LOCUS       AI288235                446 bp    mRNA    linear   EST 29-JAN-1999
DEFINITION  ql80a08.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1878614
            3' similar to contains PTR7.t2 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AI288235
VERSION     AI288235.1  GI:3931955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 796   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1878614"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-PacI;
                     Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid
                     DNA from three normalized libraries (melanocyte 2NbHM,
                     pregnant uterus NbHPU, and fetal heart NbHH19W) were
                     mixed, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from pools of 5,000 clones made from the same 3 libraries.
                     The pools consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."

  Query Match             4.5%;  Score 338.1;  DB 1;  Length 446;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  373;  Conservative    7;  Mismatches   41;  Indels    9;  Gaps    1;

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| |||||  ||  ||||||||||||||||||||||||||||||
Db        446 CTCCCGAGGAAATCTCAACTGCATGACACCTACTACGCCCCAATTCAGCAGGAAGCAGTT 387

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              |||| |||:|| |||||||||||||||||||| ||||||||||||||||||||||||||
Db        386 AGAGTGGTCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGATGGGGGAC 327

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              |||||||||||||||||||||||||||||||  |:|||||||:|||||||:|||||
Db        326 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGA 267

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||| ||| |||||||||| ||| |||||||||||||:||||||| ||||||||
```

```
Db        266 AGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCTCACACCCGACCAATCA 207

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                    |||||||| |||||||||||||||||||| ||||||||||||||||||||
Db        206 GGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGC 147

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              || ||||:||| ||:||| ||||||    || |||||||||||| |||||||||| || |
Db        146 CAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGC 87

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
              ||| |||| |||| ||| |||| ||||||||||||||| || ||||||||||||||||
Db          86 ATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTT 27

Qy       7547 TCATGCTATT 7556
              |||  |||||
Db          26 TCACTCTATT 17


RESULT 249
AQ075734/c
LOCUS       AQ075734                442 bp    DNA     linear   GSS 20-AUG-1998
DEFINITION  CIT-HSP-2359M1.TR CIT-HSP Homo sapiens genomic clone 2359M1,
            genomic survey sequence.
ACCESSION   AQ075734
VERSION     AQ075734.1  GI:3437390
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2359M1"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.5%;  Score 337.3;  DB 1;  Length 442;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  379;  Conservative   39;  Mismatches   21;  Indels    5;  Gaps    5;

Qy       2960 RTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYR 3019
              :|::|||||||||| |||||||||||||| :|||  |||||||||||| | ||:||:|||::
Db        442 GTTCTCATACCTGCACACTCTTGTCCTTTGGTACATGGATGATTTAATATTAGCTGCCCG 383

Qy       3020 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGG 3079
              |||||||||||||||| ||| ||||||||||||:|||||:|||||||||||:||||||||

Db        382 TTCAGAAACCTTGTGCAATCGAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGG 323

Qy       3080 CTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT 3139
              ||| |:|||||||||||:|:|:|||||:|||||||||||||||||||||||||||:||
Db        322 CTATAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCT 263

Qy       3140 AARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTA 3199
              |:|||||||||||||||:|||||||||||||||:::|||||||||||||||||||| |
Db        262 CAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTGA 203

Qy       3200 TCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCG 3259
              |||||||:|||||||||||||||||||| :|  |||||||||:|||:|||:|||||| |
Db        202 TCCTCATCCCAAAACCCTAAAGCAACTAAAAG-CTTCCTTGGCGTAACAGGTTTCTGCAG 144

Qy       3260 AAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAA 3319
              ||:||||||| |||||||:||||:|||||||||| :|||||:|||||: ||||||||||
Db        143 AATATGGATT-CCCAGGTATGGCGAAATAGCCAGACCATTATATACAC-AATTAAGGAAA 86

Qy       3320 CTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGG 3379
              |||||||||| |||||||| |||:|||||||||:|:||||||:::|||  ||||||||||
Db          85 CTCAGAAAGCAAATACCCA-TTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGG 27

Qy       3380 CCCCTAAAGAAGGCCTTAAACCCA 3403
              |||||||||||||| ||| || |
Db          26 -CCCTAAAGAAGGCCCTAACCCAA 4


RESULT 250
AQ876467/c
LOCUS       AQ876467                467 bp    DNA     linear   GSS 09-NOV-1999
DEFINITION  HS_2001_B2_D08_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2001 Col=16 Row=H, genomic survey
            sequence.
ACCESSION   AQ876467
VERSION     AQ876467.1  GI:6307934
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2001  row: H  column: 16
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 467.
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2001 Col=16 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
```

```
  Query Match             4.5%;  Score 337.3;  DB 1;  Length 467;
  Best Local Similarity   78.0%;  Pred. No. 0;
  Matches  366;  Conservative   41;  Mismatches   59;  Indels    3;  Gaps    2;

Qy       2947 AAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTAC 3006
             ||||||| ||:||:|::| || |||| |||||||||| |  :||| |||||||||||||
Db        466 AAGACTTCAGCCAGTTATAATNCCTGNACACTCTTGTCTTATGGTACGTGGATGATTTAC 407

Qy       3007 TTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCC 3066
             | | :  :|||::||||||||| ||||||||||||||||||||||| :|||| :|||||||
Db        406 TATAGCATGCCTGTTCAGAAACTTTGTGCCATCAAGCCACCCAAGTGCTCTAAAATTTCC 347

Qy       3067 TCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTA 3126
             ||||:|| |||||||||||:|||||||||||:|:|:|||||:|||||||| ||||||||||
Db        346 TCGCCACTTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCTCAGCAGGTTA 287

Qy       3127 AATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGC 3186
             ||||||||||:||||:|||||||||||||| |||:|||| ||| | || ||:::| |||||
Db        286 AATACTTAGGGCTAAAATTATCCAAAGGCTCCAGGGCCTTCAATAAGAAATGTAACCAGC 227

Qy       3187 CTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAA 3246
              |||||||||||||| |||||:||||||| |||||| ||||||:|::||| ||| |:|||
Db        226 TTATACTGGCTTATCTTCATCCCAAAACCTTAAAGCTACTAAGAGGGTTCTTTGACATAA 167

Qy       3247 YAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACA 3306
             :|||: |||||| ||:|||||||| | |||||: | |:|||||||||| :|||||:||||
Db        166 CAGGTATCTGCCAAATATGGATT-CTCAGGTAAGACAAAATAGCCAGATCATTATATACT 108

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
             :|||||||||||||||||| |||||||||||||||:|||||||||:| ||| ||: : |
Db        107 CTAATTAAGGAAACTCAGAAGGCCAATACCCATTTAGTAAGATGGACATCTGNAGCTGNA 48

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTT 3415
             |  | |   |||   |||     ||| | | ||| | ||: ||||||||
Db          47 GGAGATAATCAGAGACCT--NNGAGGGCCTTAACTCTAGCTCCAGTGTT 1


RESULT 251
AQ736815/c
LOCUS       AQ736815                431 bp    DNA     linear   GSS 15-JUL-1999
DEFINITION  HS_2244_B2_D10_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2244 Col=20 Row=H, genomic survey
            sequence.
ACCESSION   AQ736815
VERSION     AQ736815.1  GI:5508367
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2244  row: H  column: 20
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2244 Col=20 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.5%;  Score 336.6;  DB 1;  Length 431;
  Best Local Similarity   82.9%;  Pred. No. 0;
  Matches  355;  Conservative   37;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             | ||| |||||||||||||| : || ||| |||||| ||||||:||: ||::||||||||
Db        431 AGCTGAACACTCTTGTCCTTTGATAAGTGAATGATTAACTTTTAGCCACCCGTTCAGAAA 372

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYA---------CCTGTG 3078
             |||| |||||||||||||||||||:|||||:||||||||||:|          |  ||
Db        371 CCTTATGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCCATATTTCCTTTCCATG 312

Qy       3079 GCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRC 3138
             |||||| ||||||||||:|:|:|||||:|||||||||||||||||||||||||||||:|
Db        311 GCTACACGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGC 252

Qy       3139 TAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTT 3198
             |||: |   | || |||||||:|||||||||||||||:::|||||| ||||||||||
Db        251 TAAATTATCCAAAGGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGTCTATACTGGCTT 192

Qy       3199 ATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCC 3258
             |||||||||:|||| || ||||||||| |||:|::|||||||||:|||:|||:|||||||
Db        191 ATCCTCATCCCAAACCCTTAAAGCAACCAAGAGGGTTCCTTGGCATAACAGGCTTCTGCC 132

Qy       3259 GAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAA 3318
             |||:|||||| |||||||:||||:|||||||||||:|||||:|||||:||||||||||||
Db        131 GAATATGGATT-CCCAGGTATGGCGAAATAGCCAGGCCATTATATACACTAATTAAGGAA 73

Qy       3319 ACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAG 3378
             ||||||||||||||||||||||:|||||||||:|:||||||:::|||||||||||||||
Db          72 ACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAG 13

Qy       3379 GCCCCTAA 3386
             |||   ||
Db          12 GCCATAAA 5


RESULT 252
R68685/c
LOCUS       R68685                  463 bp    mRNA    linear   EST 26-JUN-1995
DEFINITION  yi14g06.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:139258 3', mRNA sequence.
ACCESSION   R68685
VERSION     R68685.1  GI:842202
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
```

```
            Email: est@watson.wustl.edu
            Insert Size: 1644
            High quality sequence stops: 308
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1644   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 308.
FEATURES              Location/Qualifiers
     source           1. .463
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:545824"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:139258"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares placenta Nb2HP"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.5%;  Score 336.2;  DB 1;  Length 463;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  416;  Conservative    9;  Mismatches   26;  Indels   10;  Gaps    7;

Qy       7131 CACCCCTCCTGAGGAAATCTCAGCTGCACAA--CCTCTACTACGCCCCAATTCAGCAGGA 7188
             |||||  |||| ||||| |||||||||| ||  |  ||||||  ||| |||||||| |||
Db        461 CACCCNCCCTGGGGAAANCTCAGCTGCAAAAACCTCCTACTAGCCCCNAATTCAGCNGGA 402

Qy       7189 AGCAGTTAG-AGCGGTSGTCGGCCAACCT--CCCCAACAGCACTTAGGTTTT-CCTGTTG 7244
             ||||||||  ||||||:|| |||| | |   ||||||  |||||||||||||||  |||||||
Db        401 AGCAGTTAAAAGCGGTCGTNGGCCCAACCTCCCCCAANAGCACTTAGGTTTTCCCTGTTG 342

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTT-CCTAGGCTGA--YTAAGAATCCYT 7301
             | |||||||||||||||||||||||||||||||||| ||||||||||  : |  |  ||:|
Db        341 ANATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCCTAGGCTGACCTAAGAATCCCCT 282

Qy       7302 AAGCCTAGSTGGGAAGGTGA-CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
             ||||||||:|||||||||||| |||||||||||||||||||||||||||||||||||:||
Db        281 AAGCCTAGCTGGGAAGGTGACCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 222

Qy       7361 ACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGA 7420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGA 162

Qy       7421 AATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAAC 7480
             ||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db        161 AATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAAC 102

Qy       7481 CCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 7540
             |||| |:|||||||||||||||||| ||||||||||||||||||| ||||||||||||||
Db        101 CCAANTCTTCGAGCCGGCAACGGCAAACCCCTTTGGGTCCCCTCCCNTTGTATGGGAGCT 42

Qy       7541 CTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||||||||||||||||||||||:||||
Db         41 CTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAGCTGC 1


RESULT 253
BU790503
LOCUS       BU790503                408 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in52a11.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6125493 5',
            mRNA sequence.
ACCESSION   BU790503
VERSION     BU790503.1  GI:23840987
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: in52a11.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Putative full length read
            vector to vector length is
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .408
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6125493"
                      /tissue_type="Purified pancreatic islet"
                      /lab_host="DH10B"
                      /clone_lib="HR85 islet"
                      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                      Size-selected on agarose gel. Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                      314-362-1916, Fax: 314-747-2692."

  Query Match             4.4%;  Score 332.6;  DB 1;  Length 408;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  364;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    2;

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
             | |||||||||  |||||||| ||: | ||||||||||||||||||||||  |||||||||
Db          2 CGGCAGGAAGCTGTTAGAGCAGTCATTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCT 61

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
             |||||||  |||||||||||||||||||||||||||||||||||||||| ||:|||||||:
Db         62 GTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCC 121

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
             |||||||||:|||||||||||| |||||||||||||||||||||||||||||||||:||
Db        122 TAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 181

Qy       7361 ACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGA 7420
             ||||    |||| ||||||| |||||||||||||||||||||| | ||||||| |||
Db        182 ACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTACAGA 237

Qy       7421 AATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAAC 7480
             ||||||||||||||:||||||:|||||||||||| |||||||||| |||| ||||||||||
```

```
Db        238 AATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCAGGATATAAAC 297

Qy       7481 CCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 7540
              ||| |  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        298 CCAGGCATTCGAGCTGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 357

Qy       7541 CTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||  || ||||||||||||||||||||:||||:
Db        358 CTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 397


RESULT 254
CA848433
LOCUS       CA848433                408 bp    mRNA    linear   EST 16-DEC-2002
DEFINITION  ip34e12.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6219358 5',
            mRNA sequence.
ACCESSION   CA848433
VERSION     CA848433.1  GI:26999370
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ip34e12.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Putative full length read
            vector to vector length is 409
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6219358"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

  Query Match             4.4%;  Score 332.6;  DB 1;  Length 408;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  364;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    2;

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
              | ||||||||| |||||||| ||: | ||||||||||||||||||||| |||||||||
Db          2 CGGCAGGAAGCTGTTAGAGCAGTCATTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCT 61

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
              |||||||  |||||||||||||||||||||||||||||||||||||||| ||:||||||||:
Db         62 GTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCC 121

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
              |||||||||:|||||||||||||| ||||||||||||||||||||||||||||||||:||
Db        122 TAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 181

Qy       7361 ACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGA 7420
              ||||    |||| ||||||| |||||||||||||||||||||| | |||||||| |||
Db        182 ACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTACAGA 237

Qy       7421 AATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAAC 7480
              |||||||||||||:||||||:|||||||||||| |||||||||| |||| ||||||||||
Db        238 AATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCAGGATATAAAC 297

Qy       7481 CCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 7540
              ||| |  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        298 CCAGGCATTCGAGCTGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 357

Qy       7541 CTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||  || ||||||||||||||||||||:||||:
Db        358 CTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 397


RESULT 255
AQ028930/c
LOCUS       AQ028930                429 bp    DNA     linear   GSS 30-JUN-1998
DEFINITION  CIT-HSP-2326M21.TV CIT-HSP Homo sapiens genomic clone 2326M21,
            genomic survey sequence.
ACCESSION   AQ028930
VERSION     AQ028930.1  GI:3269152
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2326M21"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.4%;  Score 332.3;  DB 1;  Length 429;
```

```
  Best Local Similarity   82.5%;  Pred. No. 0;
  Matches  363;  Conservative   37;  Mismatches   29;  Indels   11;  Gaps    3;

Qy       2975 CACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTG 3034
             || |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||| ||||
Db        429 CATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAAACCATGTG 370

Qy       3035 CCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAA 3094
             ||||||||||||||||||:|||||:||||||||||:|||||||||         |||||
Db        369 GCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC---------TCCAA 319

Qy       3095 ACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGG 3154
             || |:|:|||||:|| |||||||||||||||||||||||||:||||:||||||||||||
Db        318 ACAAAAGGCTCAGCTATGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGT 259

Qy       3155 CACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAAC 3214
             |||||:|||||||||| ||||||:::|||||||| ||||||||||||||||||:||||||
Db        258 CACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCATCCCAAAAC 199

Qy       3215 CCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCA 3274
             |||||||||||||||:|::|||||||||:|||:|| :||||||||||:||||||| ||||
Db        198 CCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGGATT-CCCA 140

Qy       3275 GGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATA 3334
             | |:  | :||||||||||:|||||: |||| ||||||||||||||||||||||||||||
Db        139 GATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATA 80

Qy       3335 CCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCC 3394
             |||||:||:|||||||||:|:||||| :::|||||||||||||| ||||||||||||||
Db         79 CCCATATAGTAAGATGGACACCTGAAACAGAAGTGGCTTTCCAGG-CCCTAAAGAAGGCC 21

Qy       3395 TTAAACCCAAGYCCCAGTGT 3414
                |||  ||  ||  :||   |||
Db         20 CTAACCCAAACCCCAGTTGT 1


RESULT 256
BF990707
LOCUS       BF990707                367 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  RC6-GN0070-271000-013-D01 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF990707
VERSION     BF990707.1  GI:12397032
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            271000-013-D01&t3=2000-10-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 359.
FEATURES    Location/Qualifiers
     source 1. .367
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="GN0070"
            /note="Organ: placenta_normal; Vector: puc18; Site_1:
            SmaI; Site_2: SmaI; A mini-library was made by cloning
            products derived from  ORESTES PCR (U.S. Letters Patent
            application No. 196,716 -  Ludwig Institute for Cancer
            Research) profiles into  the pUC 18 vector. Reverse
            transcription of tissue mRNA and cDNA amplification were
            performed under low stringency conditions."

  Query Match             4.4%;  Score 332.2;  DB 1;  Length 367;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  331;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6591 TGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGA 6650
             | |  | |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         22 TTGGCGGATCACAACCTCTACTCAGTTCTGCTACAAACTATCTCAAGAACTAAATGGGGA 81

Qy       6651 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 6710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 141

Qy       6711 AGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTG 6770
             || ||||||||:|||||||||||||||||||||:|||||||||||:|||||||||||||||
Db        142 AGCAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTG 201

Qy       6771 TTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAA 6830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 TTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAA 261

Qy       6831 AGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACC 6890
             ||||:|||||||||:||||||||||||||||:||||||:||||||||||||||||||||||
Db        262 AGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACC 321

Qy       6891 CTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTA 6936
             |||| ||||||||||||:|||||||||||||||||||||||||||||
Db        322 CTGGNGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTA 367


RESULT 257
BU790429
LOCUS       BU790429                408 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in51a12.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6125494 5',
            mRNA sequence.
ACCESSION   BU790429
VERSION     BU790429.1  GI:23840844
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: in51a12.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
```

```
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Putative full length read
            vector to vector length is
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .408
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6125494"
                      /tissue_type="Purified pancreatic islet"
                      /lab_host="DH10B"
                      /clone_lib="HR85 islet"
                      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                      Size-selected on agarose gel. Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                      314-362-1916, Fax: 314-747-2692."

  Query Match             4.4%;  Score 331;  DB 1;  Length 408;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches  363;  Conservative    9;  Mismatches   24;  Indels    6;  Gaps    2;

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
             | ||||||||| |||||||| ||: | ||||||||||||| |||||||| |||||||||
Db          2 CGGCAGGAAGCTGTTAGAGCAGTCATTGGCCAACCTCCCCATCAGCACTTGGGTTTTCCT 61

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
             |||||||  |||||||||||||||||||||||||||||||||||| ||:|||||||||:
Db         62 GTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCC 121

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
             |||||||||:|||||||||||| |||||||||||||||||||||||||||||||||:||
Db        122 TAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 181

Qy       7361 ACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGA 7420
             ||||    |||| ||||||| |||||||||||||||||||||| | |||||||| |||
Db        182 ACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTACAGA 237

Qy       7421 AATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAAC 7480
             |||||||||||||:|||||:|||||||||||| |||||||||| |||| ||||||||||
Db        238 AATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCAGGATATAAAC 297

Qy       7481 CCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 7540
             ||| |  ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        298 CCAGGCATTCGAGCTGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 357

Qy       7541 CTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||  || ||||||||||||||||||||:||||:
Db        358 CTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 397


RESULT 258
AQ242769
LOCUS       AQ242769                397 bp    DNA     linear   GSS 03-OCT-1998
DEFINITION  HS_2061_A1_B01_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2061 Col=1 Row=C, genomic survey
            sequence.
ACCESSION   AQ242769
VERSION     AQ242769.1  GI:3689343
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2061  row: C  column: 1
            Class: BAC ends
            High quality sequence stop: 397.
FEATURES              Location/Qualifiers
     source           1. .397
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=2061 Col=1 Row=C"
                      /sex="male"
                      /clone_lib="CIT Approved Human Genomic Sperm Library D"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"

  Query Match             4.4%;  Score 330.7;  DB 1;  Length 397;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  346;  Conservative   20;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       3543 TGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGT 3602
             |||||||||||||| |||:|:||| |||||: |||||||||||||||||::||||
Db          1 TGTAGTGGCAAAGGATTGGCCTCACTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGT 60

Qy       3603 ATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGT 3662
             |||||||||||||||| ||||||||:|||||||||||||||||||||||||||||||: |
Db         61 ATCTGAAGCAGTTAAAATTATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATAT 120

Qy       3663 GAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATR 3722
              ||::|||||||||||||||||||||||||||||||||||||||: ||| |||||||:
Db        121 AAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTGCTTAAATA 180

Qy       3723 TCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCC 3782
             |||||||||||||| |||||:|||||||||||| :|||| | |||| |||| ||||||||||
Db        181 TCAGGCTCTATTAATTGAAGGGCCAGTGCTGTGACTGCGTACTTTTGCAGCTCTTAACCC 240

Qy       3783 AGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTC 3842
             |  | ||||| |||||||||||||| ||| |||||:||:|||||||||||||| :| ||| |
Db        241 ACACACATTTATTCCAGACAATGAGGAANAGATAGAACATAACTGTCAACAGGTGATTGC 300

Qy       3843 TCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCA 3902
             ||| ||||| ||  ||| ||||| ||||:||||:|||||:||||||||||: | | |||
Db        301 TCANACCTACGCTGCTCAAGGGGTCCTTCTAGAGGTTCCCTTGACTGATCCCTATC-TCA 359

Qy       3903 ACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 3940
             |||||||||||||||||||||||||||||||||||||||
Db        360 ACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 397


RESULT 259
BI053386/c
LOCUS       BI053386                346 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-250101-011-E07 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI053386
VERSION     BI053386.1  GI:14460916
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            250101-011-E07&t3=2001-01-25&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 342.
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.4%;  Score 329.4;  DB 1;  Length 346;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  330;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6043 CATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGG 6102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGG 287

Qy       6103 CTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTC 6162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 CTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTC 227

Qy       6163 ARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAAC 6222
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 AGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAAC 167

Qy       6223 ACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAAC 6282
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        166 ACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAAC 107

Qy       6283 CTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGG 6342
              |||||||||||||||||||||||| | ||||||||||||||||||||||||||||||||||
Db        106 CTCACCTGTGTAAAATTTAGCAATCCCACATACACAACCAACTCCCAATGCATCAGGTGG 47

Qy       6343 GTAACTCCTCCCACACAAATAGTCTGCCTACCCTC 6377
              |||||||||||||||||||||||||||||||| |
Db         46 GTAACTCCTCCCACACAAATAGTCTGCCTACCCGC 12
```

```
RESULT 260
AQ010439
LOCUS       AQ010439                430 bp    DNA     linear   GSS 21-MAY-1998
DEFINITION  HS_2167_B2_MR_D04 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2167 Col=8 Row=H, genomic survey
            sequence.
ACCESSION   AQ010439
VERSION     AQ010439.1  GI:3143415
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2167  row: H  column: 8
            Class: BAC ends
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2167 Col=8 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.4%;  Score 329.2;  DB 1;  Length 430;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  362;  Conservative   29;  Mismatches   28;  Indels    8;  Gaps    3;

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||| ||  :| :|||||:|||||||||||| ||:|||:|||||: ||||||||||::
Db         12 AAAAGACTTGAGTAGCTCTAGGAGTCCTTACAGAGGTCCGAGGGACCAGCTTGCAACCTG 71

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::|||||:||||||||:|||||||||||||||||||:|:||| ||||||:| |
Db         72 TGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCACTGTTTATGCG 131

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
              ||||||  |||||||||||| :|||||||||||||||||||||| ||| ||||:|||||||
Db        132 TAGTGGCAGCAGTAGCAGTATTAGTATCTGAAGCAGTTAAAATGATAAAGGGAAGAGATC 191

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
              ||||  ||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db        192 TTAC-ATGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGC 250

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
              ||||||||||: | | |||||||:||||| ||||||| |||||:|||||||||||| :||
Db        251 TGTCAGACAACCTTCTGCTTAAATATCAGGTTCTATTATTTGAAGGGCCAGTGCTGTGAC 310

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
              || |||||||||||||||||||||:| ||||||||||||||||||||| ||||
Db        311 TGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAATAGAT-- 368
```

```
Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
                :|||||||||||||    |||| |||| ||||| |||||| ||||||||||:||||:|
Db        369 -GCATAACTGTCAAC----AATTGCTCATACCTACGCCACTTGAGGGGACCTTCTAGAGG 423

Qy       3878 TTCCYTT 3884
                ||||:||
Db        424 TTCCCTT 430


RESULT 261
DB375009/c
LOCUS       DB375009                357 bp    mRNA    linear   EST 03-DEC-2005
DEFINITION  DB375009 PLACE1 Homo sapiens cDNA clone PLACE1004263 3', mRNA
            sequence.
ACCESSION   DB375009
VERSION     DB375009.1  GI:83095061
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004263"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

  Query Match             4.4%;  Score 329.1;  DB 1;  Length 357;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  339;  Conservative    5;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       7227 ACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA-TTTCCTAGG 7285
              | |  | ||||||||||||||||||||||||||||||||||||||| |||| |||| |||
Db        356 ATTAGGTTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGNTGGATTTTCNTAGG 297

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
              | || ||||||||| ||||| ||| |||||||||||||||||||||||||||||||||||
Db        296 CNGANTAAGAATCCNTAAGCNTAGNTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 237

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 7405
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 177

Qy       7406 GACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAAT 7465
              |||||||||||||||||||||||||||:||||||:||||||||||||||||||||||||

Db        176 GACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACAGCAGGAGGGACAAT 117

Qy       7466 GATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCC 7525
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        116 GATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCC 57

Qy       7526 CTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          56 CTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 262
AQ225099/c
LOCUS       AQ225099                461 bp    DNA     linear   GSS 20-SEP-1998
DEFINITION  HS_2017_A2_H10_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2017 Col=20 Row=O, genomic survey
            sequence.
ACCESSION   AQ225099
VERSION     AQ225099.1  GI:3640942
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2017  row: O  column: 20
            Class: BAC ends
            High quality sequence stop: 461.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2017 Col=20 Row=O"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.4%;  Score 326.7;  DB 1;  Length 461;
  Best Local Similarity   81.3%;  Pred. No. 0;
  Matches  347;  Conservative   38;  Mismatches   39;  Indels    3;  Gaps    2;

Qy       2989 RGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCC 3048
              :|||:|||||  ||||||||||:||:||| :||||||||| |||||||||| | |||| |
Db        424 GGTATGTGGACNATTTACTTTTAGCTGCCAGTTCAGAAACTTTGTGCCATCGAACCACNC 365

Qy       3049 AAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARC 3108
              ||| :|||||:|| ||||| ||: ||||||||||||:||||||||||:|:|:|||||:|
Db        364 AAGTGCTCTTAAACTTCCTTGCTGCCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGC 305

Qy       3109 TCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
              |||||| | |||||||||||||||||: || :|||||||||||||||||:|||||||||
Db        304 TCTGCTTATAGCAGGTTAAATACTTAGGGTTAGAATTATCCAAAGGCACCAGGGCCCTCA 245

Qy       3169 GTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAA 3228
              |||||| |:::|| ||||||||||||||||||||||:|||| |||| |||||| |||
```

```
Db        244 GTGAGGCACATATTCAGCCTATACTGGCTTATCCTCATCCCAAAGCCCTGAAGCAATTAA 185

Qy       3229 GRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATA 3288
              |:|: |||||||||:|||:|||:||||||| ||:||||||| |||||||: |||:|||||
Db        184 GAGGCTTCCTTGGCATAACAGGCTTCTGCCAAATATGGATT-CCCAGGTACGGCGAAATA 126

Qy       3289 GCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGA 3348
              ||||||:|||||: ||||:||||| |||||||||||||||||||| ||||||:||  |
Db        125 GCCAGGCCATTATGTACACTAATTTAGGAAACTCAGAAAGCCAATACTCATTTAGTAGAA 66

Qy       3349 TGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
              ||||:|:||||||:::|||||||||| ||||||   ||  | | | |  ||| ||||:||
Db         65 TGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCATAAAGAAGGCCAT--AACNCAAGCCC 8

Qy       3409 CAGTGTT 3415
              |||| ||
Db           7 CAGTATT 1
```

```
RESULT 263
T69704/c
LOCUS       T69704                363 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  yd13a03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:66988 3', mRNA sequence.
ACCESSION   T69704
VERSION     T69704.1  GI:680852
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 766
            High qality sequence stops: 341 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 766   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 341.
FEATURES             Location/Qualifiers
     source          1. .363
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:463693"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66988"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
```

```
  Query Match             4.4%;  Score 326.4;  DB 1;  Length 363;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  345;  Conservative    9;  Mismatches    8;  Indels    2;  Gaps    2;

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 304

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              ||||||||||:|||||||||:|| ||||||:||||||||||||||||||||||||||||
Db        303 TCCTAGGCTGACTAAGAATCCCTAGGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAA 244

Qy       7339 CACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATT 7398
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        243 CACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATT 184

Qy       7399 AGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAG 7458
              ||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||
Db        183 AGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAG 124

Qy       7459 GGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGT 7518
              |||||||||||||||||||||||||||||:|||||  ||||||||||| ||||||||||
Db        123 GGACAATGATCGGGATATAAACCCAAGTCTTCGA-NCGGCAACGGCAAACCCCTTTGGG- 66

Qy       7519 CCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARC 7578
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  |:
Db         65 CCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTNAAAG 6

Qy       7579 TGCR 7582
                |:
Db           5 NACA 2
```

```
RESULT 264
CA432887/c
LOCUS       CA432887              438 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-CO0-aqv-a-12-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
            UI-H-CO0-aqv-a-12-0-UI 3', mRNA sequence.
ACCESSION   CA432887
VERSION     CA432887.1  GI:24797307
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 28-351, >LTR17#LTR/Retroviral (matched compliment)
            352-438, >HERV17#LTR/Retroviral (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-CO0-aqv-a-12-0-UI"
                     /tissue_type="mixed"
                     /dev_stage="mixed"
```

/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Sub9"
/note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; tissues: Cholonic mucosa with Crohns disease, Cholonic mucosa with ulcerative colitis, Fetal thymus, Cervix, Cervical adenosquamous carcinoma, Ligament cells, Prostate carcinoma, Bladder carcinoma, Brain oligodenroga ; NCI_CGAP_Sub9 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are CGTC, AACG, GGGCC, GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_TISSUE=Prostate Carcinoma
TAG_LIB=UI-H-CO0
TAG_SEQ=ATGG"

Query Match 4.3%; Score 324.7; DB 1; Length 438;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 366; Conservative 9; Mismatches 27; Indels 21; Gaps 2;

```
Qy       7169 TACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCAC 7228
             || ||||||| |||||||||||||||||||||||: | |||| ||||||||||||||||
Db        438 TATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATTGGCCCACCTCCCCAACAGCAC 379

Qy       7229 TTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTG 7288
             || |||||||||||||||| |||||||||| |||||||||||||||||||||| ||||
Db        378 TTGGGTTTTCCTGTTGAGAAGGGGGACTGAAAGACAGGACTAGCTGGATTTCCCAGGCCA 319

Qy       7289 AYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTT 7348
             |:||||||||:|||||||||:|||||||||||||| |||||||||||||||| ||||||
Db        318 ACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACAGGGCTT 259

Qy       7349 GCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTA 7399
             |||||||||:||||||| ||||||||         |||||||||||||||||||||||||
Db        258 GCAACTTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTA 199

Qy       7400 GGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGG 7459
             |||||| || |||||||||||||||||||||:||||||:|||||||||||  |||||
Db        198 GGCAAAAACTGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGTGGGAGG 139

Qy       7460 GACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTC 7519
             |||| |||||||||||||||| |  ||||||| ||||| ||||||||||||||||||||
Db        138 GACAGTGATCGGGATATAAACCCAGGCATTCGAGCTGGCAATGGCAACCCCCTTTGGGTC 79

Qy       7520 CCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCT 7579
             ||||||| || |||||||||||||||||||            || ||||||||||:||
Db          78 CCCTCCCATTTTATGGGAGCTCTGTTTTCA-----------CTCTTAAATCTTGCAACT 31

Qy       7580 GCR 7582
             ||:
Db          30 GCA 28
```

RESULT 265
BQ365139
LOCUS BQ365139 354 bp mRNA linear EST 21-MAY-2002
DEFINITION RC6-GN0070-170800-011-D03 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ365139
VERSION BQ365139.1 GI:21040651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 354)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-170800-011-D03&t3=2000-08-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 354.
FEATURES Location/Qualifiers
source 1. .354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0070"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 4.3%; Score 323.9; DB 1; Length 354;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 332; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```
Qy       6591 TGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGA 6650
             | |  | |||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          10 TTGGCGGATCACAACCTCTACTCAGTTCTACT-CAAACTATCTCAAGAACTAAATGGGGA 68

Qy       6651 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 6710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          69 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 128

Qy       6711 AGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTG 6770
             ||||||||||:|||||||||||||||||||:||||||||||||:|||||||||||||||
Db         129 AGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTG 188

Qy       6771 TTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAA 6830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 TTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAA 248

Qy       6831 AGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACC 6890
             ||||:|||||||||:|||||||||||||||||:||||||:||||||||||||||||||||
Db         249 AGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACC 308

Qy       6891 CTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTA 6936
             ||| |||||||||||||:||||||||||||||||||||||||||||
Db         309 CTGNGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTA 354
```

RESULT 266
B95002
LOCUS B95002 418 bp DNA linear GSS 26-JUN-1998

```
DEFINITION  CIT-HSP-2167H8.TR CIT-HSP Homo sapiens genomic clone 2167H8,
            genomic survey sequence.
ACCESSION   B95002
VERSION     B95002.1  GI:2977339
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2167H8.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .418
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7101930"
                     /db_xref="taxon:9606"
                     /clone="2167H8"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.3%;  Score 323.7;  DB 1;  Length 418;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  364;  Conservative   30;  Mismatches   24;  Indels    9;  Gaps    4;

Qy       3430 GGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACA 3489
              ||||||||||||:||:|||::|||||| |||| ||||:| :|||||:|||||||||||||
Db             1 GGCAAGACTTTTCTTTATATGTCACAG-AAAAGACAGGAGTAGCTCTAGGAGTCCTTACA 59

Qy       3490 CAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTG 3549
              ||:|||:|||||: |||||||||||::||||::|||||:||||||||:|||||||||||
Db            60 GAGGTCCGAGGGACCAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTAGTG 119

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||||||:|:||| ||||||:| |||||||  |||||||||| :||||||||||||
Db           120 GCAAAGGGTTGGCCTCACTGTTTATGCGTAGTGGCAGCAGTAGCAGTATTAGTATCTGAA 179

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |||||||||||  ||| ||||:|||||||||||  |||||||||||||||||:|||||::||
Db           180 GCAGTTAAAATGATAAAGGGAAGAGATCTTAC-ATGTGGACATCTCATGATGTGAACGGC 238

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              |||||||||||||||||||||||||||||||||||||: | | |||||||:|||||| |
Db           239 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCTTCTGCTTAAATATCAGGTT 298

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              |||||| |||||:|||||||||| :|||| |||||||||||||||||||||:| ||
Db           299 CTATTATTTGAAGGGCCAGTGCTGTGACTGCGCACTTGTGCAACTCTTAACCCAGCCACA 358

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              |||||||||||||||||||||||||    :||||||||||||    |||  ||||||||
Db           359 TTTCTTCCAGACAATGAAGAAAAGAT---GCATAACTGTCAAC----AATAGCTCAAACC 411

Qy       3850 TATGCCA 3856
              || ||||
Db           412 TACGCCA 418


RESULT 267
AA860368/c
LOCUS       AA860368                342 bp    mRNA    linear   EST 31-DEC-1998
DEFINITION  aj59c05.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1394600
            3' similar to contains PTR7.t1 PTR5 repetitive element ;, mRNA
            sequence.
ACCESSION   AA860368
VERSION     AA860368.1  GI:2954363
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1305   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1394600"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares_testis_NHT"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from mRNA obtained from
                     Clontech Laboratories, Inc., and primed with  a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "

  Query Match             4.3%;  Score 323;  DB 1;  Length 342;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  320;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7250 GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAG 7309
              ||||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||
Db           342 GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAG 283

Qy       7310 STGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGAC 7369
              :||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db           282 CTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGAC 223

Qy       7370 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAA 7429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           222 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAA 163
```

```
Qy       7430 TCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTT 7489
              ||||:||||||:||||||||||||||||||||||||||||||||||||||||||||:||
Db        162 TCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTT 103

Qy       7490 CGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              ||   |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        102 CGGAGCGGCAACGGCAAACCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 43

Qy       7550 TGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              ||||||||||||||||||||||||||:||||:
Db          42 TGCTATTTCACTCTATTAAATCTTGCAACTGCA 10


RESULT 268
BQ365143/c
LOCUS       BQ365143                346 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC6-GN0070-170800-011-F12 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365143
VERSION     BQ365143.1  GI:21040655
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            170800-011-F12&t3=2000-08-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.3%;  Score 321.6;  DB 1;  Length 346;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches  324;  Conservative    8;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
              |||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||

Db        346 TGGCATTGGCGGTATCGCAACCTCTACTCAGTTCTACTACAAACTATCTCGAGAACTAAA 287

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 227

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGG 6764
              ||||||||||||||||:||||||||||||||||||:||||||||||||:|||||||||
Db        226 AGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGG 167

Qy       6765 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 6824
              |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 AGTCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 107

Qy       6825 TGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACAC 6884
              ||||||||||:|||||||||:||||||||||||||||:||||||:|||||||||||||||
Db        106 TGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACAC 47

Qy       6885 TGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
              |||||||||||||||||||||||:|||||  |  |   |  |||
Db          46 TGGACCCTGGGGCCTCCTCAGCCAATGGACGTCTGGACTAGTCC 3


RESULT 269
AQ345621
LOCUS       AQ345621                430 bp    DNA     linear   GSS 07-MAY-1999
DEFINITION  RPCI11-116N15.TV RPCI-11 Homo sapiens genomic clone RPCI-11-116N15,
            genomic survey sequence.
ACCESSION   AQ345621
VERSION     AQ345621.1  GI:4170517
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-116N15.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7544486"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-116N15"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             4.3%;  Score 320.9;  DB 1;  Length 430;
  Best Local Similarity   85.4%;  Pred. No. 0;
```

```
  Matches  368;  Conservative   15;  Mismatches   39;  Indels    9;  Gaps    3;

Qy       4161 CYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGA 4220
              |:|||:|||||||||| ||:||| |||||||||||||||:|||||||:||||:|||||||
Db          1 CCGAGGGAACACCTAACAAACACCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGA 60

Qy       4221 AACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCA-----NAAAGGAAAGRAAAG 4275
              ||||||:||||||||:| ||| |||||||:||||||| ||      ||||||||:|||| 
Db         61 AACCTAAAGAGGTGGCAATCTAACACTGCCGGGGTCACCAGAAAGGAAAGGAAAGGAAAG 120

Qy       4276 GGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCC 4335
              |||||||:|||:|||:|||||||||:|||||||| | |||||||| ||||||  | ||||
Db         121 GGAAATAGAAGGGAATTGCCAAGCAGATATTGAACCGAAAAGAGCCGCAAGGTGGAACCC 180

Qy       4336 TCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAA 4395
              |||||||||||| |||||  ||   ||| |||||  | |||||||| || ||||||||||
Db         181 TCCATTAGAAATACTTATAGAAGGACCCCTAGTAAGGTGTAATCCCCTCTGGGAAACCAA 240

Qy       4396 GCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCC 4452
              |||||||||||||||||| ||||| |||||||||||||||||| |||    ||||| |||||
Db         241 GCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCACAAGGACATAGTTTCCTCCC 300

Qy       4453 CTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAA 4512
              ||| ||| || |||||||||||||| |||||||||||||||||||| ||| |||||||||
Db         301 CTCAGGATGGCTAGCCACTGAAGAA-GGAAAATACTTTTGCCTGCAGCTAACCAATGGAA 359

Qy       4513 ATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGC 4572
              ||||||||||| |||||  |||||||||||||||| |||| |||||||||||||:|||||
Db         360 ATTACTTAAAATTCTTCACCAAACCTTTCACTTANGCATNGATAGCACCCATCAGATGGC 419

Qy       4573 CAAATCATTAT 4583
              |||||||||||
Db         420 CAAATCATTAT 430


RESULT 270
T47345/c
LOCUS       T47345                342 bp    mRNA    linear   EST 01-FEB-1995
DEFINITION  yb10h02.s1 Stratagene placenta (#937225) Homo sapiens cDNA clone
            IMAGE:70803 3', mRNA sequence.
ACCESSION   T47345
VERSION     T47345.1  GI:649326
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 598
            Source: IMAGE Consortium, LLNL This clone is available royalty-free
            through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov)
            for further information.
            Insert Length: 598   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 342.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:491700"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:70803"
                     /sex="male"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene placenta (#937225)"
                     /note="Organ: placenta; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP
                     XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

  Query Match             4.3%;  Score 320.8;  DB 1;  Length 342;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  319;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||
Db          342 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTA 283

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||:|||||||||| ||||||||||||||||||||||||||||||||||:||||
Db          282 AGCCTAGCTGGGAAGGTGNCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCAC 223

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          222 ACCTGNCCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 163

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
              ||||||||||:||||||:||||||||||||||||||||||||||||||||||||||||||
Db          162 TAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 103

Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
              ||||:||||   |||||||||||| ||||||||||||||||||||||||||||||||||
Db          102 AAGTCTTCGGAGCGGCAACGGCAANCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 43

Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||||||||||||||      |:||||:
Db           42 GTTTTCATGCTATTTCACTCTATTAAAANNAANAACTGCA 3


RESULT 271
AQ200416
LOCUS       AQ200416                416 bp    DNA     linear   GSS 20-APR-1999
DEFINITION  RPCI11-48P23.TK RPCI-11 Homo sapiens genomic clone RPCI-11-48P23,
            genomic survey sequence.
ACCESSION   AQ200416
VERSION     AQ200416.1  GI:3612615
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
```

```
           Research Genetics (info@resgen.com). BAC end search page:
           http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
           Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7518430"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-48P23"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             4.3%;  Score 320.6;  DB 1;  Length 416;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  359;  Conservative   15;  Mismatches   39;  Indels    8;  Gaps    2;

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
Db          1 CCTAACTTCCGAGGAAACACCTATCAAACATCAGGAAGCCATTAGGAAATTAGTATTGGC 60

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
Db         61 TGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCTGGGTCATCA-----GAAAGGA 115

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
Db        116 AAGGGAAATAAAAGAGAACCACCAAGCAGATATTAAAGCAAAAAGAGCCACAAGGCAGGA 175

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
Db        176 TGCTCCATTAGAAATGCTTACAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAAC 235

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
Db        236 AAAACCCCAATATTCAGCAGAAGAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCT 295

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
Db        296 CCCCTCAGGATGGCTAGCCACCGAAGAGGGAAAAATACTTTTGCCTGCAGCTAACCAATG 355

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
Db        356 GAAATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGAT 415

Qy       4570 G 4570
Db         416 G 416


RESULT 272
AQ701255/c
LOCUS       AQ701255                449 bp    DNA     linear   GSS 07-JUL-1999
DEFINITION  HS_2132_A2_B08_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2132 Col=16 Row=C, genomic survey
            sequence.
ACCESSION   AQ701255
VERSION     AQ701255.1  GI:5410681
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2132  row: C  column: 16
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2132 Col=16 Row=C"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.3%;  Score 320.6;  DB 1;  Length 449;
  Best Local Similarity   80.5%;  Pred. No. 0;
  Matches  352;  Conservative   39;  Mismatches   42;  Indels    4;  Gaps    3;

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
Db        441 CCAAGACTTGAGCCAGTTCT--TTACTTGACACTCATGTCCTTTGGTACATGGATGA-TT 385

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
Db        384 ACTTTAAGCTGCCGATTCAGAACCGTTGTGCCATCAAGCCACCCAAGCGCTTTTATACTT 325

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
Db        324 CCTCGCCACTTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGT 265

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
Db        264 TAAACACTTACGGCTAAAATTATCCANAGGCACCAGGGCCNTCAGTGAGAACGGTATCCA 205

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
Db        204 GCCTATATCGGCTTATCCTCATCCCAAACGCNTGAAGCAATTAAGAATGTTCCTTGGCAT 145

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
Db        144 AACAGGCTTCTGCCAAATATGGATT-AACAGGTATGGCAAATAAGCCAGGCCATTATACA 86

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
Db         85 CACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATAGACACCTGAAGCAG 26

Qy       3365 AAGTGGCTTTCCAGGCC 3381
Db         25 AAGTGGCTNTTCAGGTC 9


RESULT 273
AQ150723
LOCUS       AQ150723                422 bp    DNA     linear   GSS 08-OCT-1998
DEFINITION  HS_3203_A2_F07_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3203 Col=14 Row=K, genomic survey
            sequence.
ACCESSION   AQ150723
VERSION     AQ150723.1  GI:3544181
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3203  row: K  column: 14
            Class: BAC ends
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate:3203 Col=14 Row=K"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.2%;  Score 317.3;  DB 1;  Length 422;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  363;  Conservative    6;  Mismatches   41;  Indels   13;  Gaps    2;

Qy       6793 TGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGA 6852
             |||| |||||||||||||||| ||||| || || |||||| ||:|||||||| |||||||
Db          1 TGTTGTTATGTTAATCAATCTGGAATTGTTACCGAGAAATTTAAAGAAATTCAAGATCGA 60

Qy       6853 ATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGG 6912
             || ||| |||:|||||||:||||||| |||||||||||||| |||||||||||||:||||
Db         61 ATGCAATGTAGAGCAGAGGAGCTTCAAAACACTGGACCCTGNGGCCTCCTCAGCCAATGG 120

Qy       6913 ATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTT 6972
             |||||||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        121 ATGCCCTGAATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTT 180

Qy       6973 GGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAA 7032
             |||||||||||||| |  |||||||||| |||||||||||||||||| |||||||||:||
Db        181 GGACCCTGTATCTTTTATTTCCTTGTTAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAAG 240

Qy       7033 CTACAAA----------TGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             || ||||           |||||||| |||||||||||| |||||| |||||||   ||
Db        241 CTGCAAATCGTTCTTAAATTGGAGCCCCAGATGCAGTCCATGACTAANATCTACCAAGGA 300

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||| ||||| |||||||:||||||| | || | |||||||||||  ||||||||||||
Db        301 CCCTTGGACTGGCCTGCTAGCCCATGCTCCG-TGTTAATGACATTGAAGGCACCCCTCCC 359

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGC 7200
              |||||||||| |||||||  | ||| || ||||||||||||||||||||||||||||||
Db        360 AAGGAAATCTCAACTGCACANACCCTATTATGCCCCAATTCAGCAGGAAGCAGTTAGAGC 419

Qy       7201 GGT 7203
              |||
Db        420 GGT 422


RESULT 274
BU790130/c
```

```
LOCUS       BU790130                404 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in51a12.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6125494 3',
            mRNA sequence.
ACCESSION   BU790130
VERSION     BU790130.1  GI:23840267
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .404
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6125494"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

  Query Match             4.2%;  Score 317;  DB 1;  Length 404;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  355;  Conservative    8;  Mismatches   33;  Indels    6;  Gaps    2;

Qy       7179 TTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 7238
             | | ||||||||| |||||||| ||: | |||||||||||||||||||||||| |||||||
Db        396 TCCGGCAGGAAGCTGTTAGAGCAGTCATTGGCCAACCTCCCCAACAGCACTTGGGTTTTC 337

Qy       7239 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATC 7298
             |||||||| |||||||||||||||||||||||||||||||||||| ||:||||||||
Db        336 CTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATC 277

Qy       7299 CYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGY 7358
             |:|||||||||:|||||||||||||| |||||||||||||||||||||||| ||    | :
Db        276 CCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGAAAAAAAAC 217

Qy       7359 TCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAA 7418
             ||||||    |||| ||||||| |||||||||||||||||||||| | |||||||| |
Db        216 TCACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTACA 161

Qy       7419 GAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAA 7478
```

```
              |||||||||||||||:||||||:||||||||||| |||||||||| |||| ||||||||
Db        160 GAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCAGGATATAA 101

Qy       7479 ACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAG 7538
              ||||| |  ||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        100 ACCCAGGCATTCGAGCTGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAG 41

Qy       7539 CTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              ||||||||||  || |||||   ||| ||||||||||:|||
Db         40 CTCTGTTTTCA--CTGTTTCAGGGTATAAAATCTTGCAACTG 1


RESULT 275
DB352776/c
LOCUS       DB352776                420 bp    mRNA    linear   EST 04-DEC-2005
DEFINITION  DB352776 UTERU3 Homo sapiens cDNA clone UTERU3014327 3', mRNA
            sequence.
ACCESSION   DB352776
VERSION     DB352776.1  GI:83191986
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 420)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .420
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UTERU3014327"
                     /tissue_type="uterus"
                     /clone_lib="UTERU3"
                     /note="Vector: pME18SFL3"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 420;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  348;  Conservative    7;  Mismatches   40;  Indels    9;  Gaps    1;

Qy       7162 CCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCA 7221
              |  ||  ||||||||||||||||||||||||||||||| |||:|| ||||||||||||||
Db        420 CNCCTNNTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTCGTTGGCCAACCTCCCCA 361

Qy       7222 ACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCC 7281
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ACAGCACTTGGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCC 301

Qy       7282 TAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACAC 7341
              |||||  |:|||||||||:|||||||||:|||||||||||| ||| || ||||||| |||
Db        300 TAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTACACCCCCCTTTAACCAC 241

Qy       7342 GGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATG 7392
              |||||||||||||:|||||| ||||||||          ||||||||  ||||||||
Db        240 TAGGCTTGCAACTTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATG 181

Qy       7393 CTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAG 7452
              |||||||||||||| ||||||||||||||||||||||| ||||:||| ||:||| ||||||
Db        180 CTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAA 121

Qy       7453 CAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCT 7512
              || |||||||||||| |||||||||| || |  ||| |||| |||| ||| |||| ||
Db        120 GGGGCGGGACAATGATCAGGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACT 61

Qy       7513 TTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATT 7556
              |||||||||||||| || ||||||||||||||||||  |||||
Db         60 TTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCACTCTATT 17


RESULT 276
BF761543
LOCUS       BF761543                421 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  CM0-CS0041-051000-592-c12 CS0041 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF761543
VERSION     BF761543.1  GI:12109443
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM0&t2=CM0-CS0041-
            051000-592-c12&t3=2000-10-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CS0041"
                     /note="Organ: colon_est; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
```

```
  Query Match             4.2%;  Score 314.7;  DB 1;  Length 421;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  339;  Conservative   14;  Mismatches   37;  Indels    1;  Gaps    1;

Qy       4431 CCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTT 4490
             | | | ||   ||||| |||||||| ||| || ||||||| |||||||| ||||||||||
Db         31 CATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGAAAAATACTTT 90

Qy       4491 TGCCT-GCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGC 4549
             ||||| ||| ||| ||||||||||||||||||||||||||| || ||||||||||||||||
Db         91 TGCCTGGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAGACCTTTCACTTAGGC 150

Qy       4550 ATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACT 4609
             || |||||||||||||:||||||||||| |||||||||||||||| ||||||||||||
Db        151 ATTGATAGCACCCATCAGATGGCCAAATCGTTATTTACTGGACCAGGACTTTTCAAAACT 210

Qy       4610 ATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGC 4669
             ||||||||:|||:|||||||||||||||:|||||||:|:|||||||||||||||||:|||||
Db        211 ATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAAAGAAATAATCCCCTGCCTTATCGC 270

Qy       4670 CAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATT 4729
             ||||||||||||||:||||||:|||||||||||||||  |: :| :   |  |   ||||
Db        271 CAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCAGGAGAACACTGGCAACTAGATT 330

Qy       4730 TTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCA 4789
             ||||||||| ||||||| |||||||||||||||||||||||||||||||||:||||| |||
Db        331 TTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTCTCA 390

Qy       4790 CGGGTTGGGCARAGGCCTTCCCCTGTAGGAC 4820
             | |||||||| :|||||||||| ||||||||
Db        391 CTGGTTGGGCGGAGGCCTTCCCTTGTAGGAC 421


RESULT 277
AQ005223
LOCUS       AQ005223                399 bp    DNA     linear   GSS 27-JUN-1998
DEFINITION  CIT-HSP-2290E6.TR CIT-HSP Homo sapiens genomic clone 2290E6,
            genomic survey sequence.
ACCESSION   AQ005223
VERSION     AQ005223.1  GI:3082668
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2290E6.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7150593"
                     /db_xref="taxon:9606"
                     /clone="2290E6"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.2%;  Score 313.6;  DB 1;  Length 399;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  339;  Conservative   32;  Mismatches   22;  Indels    4;  Gaps    3;

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG-AAAAAAACAGRAAYAG 3472
             | |||:||||||||:| ||||||||||:||:|||::|||||| ||||||| ||:||:||
Db           5 TCAAGCTTGCCAACGGTGCAAGACTTTTCTTTATATGTCACAGAAAAAAAAATAGGAATAG 64

Qy       3473 CTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTA 3532
             ||||:||||||||||||||:|| :|||||: ||||||||||::|||||::|| |||:||
Db          65 CTCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCTGTGGCATACTTGAGTA 124

Qy       3533 AGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAG 3592
             || |||:|||||||||||||||||||||:|:|||||||||: ||||||||  |||||||
Db         125 AGAAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCAGCAGTAG 184

Qy       3593 CAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTA-CTGTGTGGACA 3651
             ||||::|||  || |||||||||||||||||||| ||:|||||||||  |||||||||
Db         185 CAGTCTTAGCCTCCGAAGCAGTTAAAATAATACAAGGAAGAGATCTTATGTGTGTGGACA 244

Qy       3652 TCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGT 3711
             ||||||||:|||||::|||||||||||||||||||||||||||||||||||||||||||: |
Db         245 TCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCAT 304

Qy       3712 TTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCA 3771
             |  |||||||:|||||||  ||| |||||:|||||||||||:|||| ||||||||||
Db         305 TAGCTTAAATATCAGGCTC--TTATTTGAAGGGCCAGTGCTGCGACTGCGCACTTGTGCA 362

Qy       3772 ACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAG 3808
             ||||||||||||:| | ||||||||||||||||||
Db         363 ACTCTTAACCCAGCCACGTTTCTTCCAGACAATGAAG 399


RESULT 278
AQ062249/c
LOCUS       AQ062249                385 bp    DNA     linear   GSS 31-JUL-1998
DEFINITION  CIT-HSP-2347N16.TR CIT-HSP Homo sapiens genomic clone 2347N16,
            genomic survey sequence.
ACCESSION   AQ062249
VERSION     AQ062249.1  GI:3364161
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2347N16.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
```

```
     source          1. .385
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2347N16"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.2%;  Score 312.1;  DB 1;  Length 385;
  Best Local Similarity   87.8%;  Pred. No. 0;
  Matches  338;  Conservative   10;  Mismatches   34;  Indels    3;  Gaps    1;

Qy       6651 CATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGC 6710
             ||||||| ||||| | |||| ||||||| || |||||||| || ||||||||||||||||||
Db        385 CATGGAATGGGTCTCTGACTTCCTGGTCGCCGTGCAAGACCACCTTAACTCCCTAGCAGC 326

Qy       6711 AGTAGTCCTTCRAAATCGAAGAGCTTTAGACTY---GCTAACCGCTGARAGAGGGGGAAC 6767
             ||||||||||:|||||||||||||||||||:   ||||| |||  |:||||||||||
Db        325 AGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTTAGGCTAATCGCCAAAAGAGGGGGAAC 266

Qy       6768 CTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGA 6827
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| ||
Db        265 CTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCAGAATCGTCACCGA 206

Qy       6828 GAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGG 6887
             || ||||:|||||||||:|||| | ||||||| ||:||||||:||||||| |||||||||
Db        205 GACAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAGAGGAGCTTCAAAACACTGG 146

Qy       6888 ACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGC 6947
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        145 ACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGC 86

Qy       6948 AGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGT 7007
             |||||||||||||| ||||||||||||||||||||||||||:|||||||||||||   |||
Db          85 AGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGCTTGA 26

Qy       7008 CTCTTCCAGAATCGAAGCTGTRAAA 7032
               | ||| | | |   | ||  :| |
Db          25 GTATTCTATAGTGTCACCTAAGATA 1


RESULT 279
AQ044774
LOCUS       AQ044774                377 bp    DNA     linear   GSS 14-JUL-1998
DEFINITION  CIT-HSP-2319K15.TR CIT-HSP Homo sapiens genomic clone 2319K15,
            genomic survey sequence.
ACCESSION   AQ044774
VERSION     AQ044774.1  GI:3313701
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 377)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2319K15.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .377
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2319K15"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             4.2%;  Score 312;  DB 1;  Length 377;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  330;  Conservative   25;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::||||||:||||||||:|||||||||||||||||||:|:|| |||||||:||
Db             1 GTGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCCTTGTTTATGG 60

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||||||| |   ||||||||::|||||||||||||||||||||| |||||||:||||||
Db            61 GTAGTGGCGTTCGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGAT 120

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              |||||||||||||||||||||:|||||::|||||  || ||||||||||||||||||||
Db           121 CTTACTGTGTGGACATCTCATGATGTGAACAGCATATTCTCTGCTAAAGGAGACTTGTGG 180

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAG-GCTCTATTACTTGAARGGCCAGTGCTGCR 3755
              |||||||||| :||||||||||:||||     |||||||||||:||||||| || :
Db           181 CTGTCAGACAATCGTTTACTTAAATATCAGTAAAATATTACTTGAAGGGCCAGTGTTGTG 240

Qy       3756 ACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGAT 3815
              |||||||||||||||||||||||||||||:| | ||||||||||||||||||||||||
Db           241 ACTGTGCACTTGTGCAACTCTTAACCCAGCCACGTTTCTTCCAGACAATGAAGAAAAGAT 300

Qy       3816 ARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGA 3875
              |:||:|||||||||||||:|||||  |||||||||  |||||  |||||  ||||:||||
Db           301 AGAACATAACTGTCAACAAATAATTGCTCAAACCTACTCCACTTGAGGGG-CCTTTTAGA 359

Qy       3876 RGTTCCYTTGACTGATCC 3893
              :|||||:||||||||||
Db           360 GGTTCCCTTGACTGATCC 377


RESULT 280
AQ040947
LOCUS       AQ040947                413 bp    DNA     linear   GSS 14-JUL-1998
DEFINITION  CIT-HSP-2326C16.TV CIT-HSP Homo sapiens genomic clone 2326C16,
            genomic survey sequence.
ACCESSION   AQ040947
VERSION     AQ040947.1  GI:3310218
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2326C16.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
```

```
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .413
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="2326C16"
                      /sex="Male"
                      /cell_type="Sperm"
                      /clone_lib="CIT-HSP"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII"

  Query Match             4.2%;  Score 311.9;  DB 1;  Length 413;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  368;  Conservative    1;  Mismatches   42;  Indels    3;  Gaps    3;

Qy       1616 AAGTTTATCACTCAGTCAGCTGCAGACATTAGAA-AAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||| |||||||||||||||||||||||||||||| ||||||||||||||| | | || ||
Db          2 AAGCTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGTCTTATGC 61

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              || ||||||||||||||||||||||| ||||||| |||:| ||||||||||||||||||||
Db         62 CCCGAGCAAAACTTAGAAACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGAT 121

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||| |||||||||||||| |||| ||| ||| |||||| ||||||||||||||||||||
Db        122 CAGGAAGAGCAGGCGGAACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCAT 180

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              | ||||| |||||| |||||||||||||||||||| ||||||| ||||||||||    ||
Db        181 GGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACAT 239

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
              |||||||||||||||||| || |||||||||| ||||||||||||||||||||||||||
Db        240 GCCTAATAGGGCTTGCTTTCACTGCGGTCTACACGGACACTTTAAAAAAGATTGTCCAAA 299

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
              ||||| |||||| ||    | || |||||||||||| | ||||||||||||||||||||||
Db        300 TAGAAATAAGCCACCTTCTTGTTCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCA 359

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCA 2028
              ||  ||||||| |  |||||| | || |||||||||||||||||||||||||||
Db        360 CTTTCCCAGGGAATGAAGGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCA 413


RESULT 281
AA577847/c
LOCUS       AA577847                427 bp    mRNA    linear   EST 12-SEP-1997
DEFINITION  nn24g11.s1 NCI_CGAP_Gas1 Homo sapiens cDNA clone IMAGE:1084868 3'
            similar to TR:G38333 G38333 ENDOGENOUS RETROVIRUS PHE.1 PROTEIN. ;,
            mRNA sequence.
ACCESSION   AA577847
VERSION     AA577847.1  GI:2356031
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 3854   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 371.
FEATURES              Location/Qualifiers
     source           1. .427
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1084868"
                      /tissue_type="gastric tumor"
                      /lab_host="SOLR (kanamycin resistant)"
                      /clone_lib="NCI_CGAP_Gas1"
                      /note="Organ: stomach; Vector: Bluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                      Oligo dT. Pooled gastric tumors. 5' adaptor sequence: 5'
                      GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                      CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size:  1.0 kb."

  Query Match             4.2%;  Score 311.9;  DB 1;  Length 427;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  361;  Conservative   37;  Mismatches   28;  Indels    9;  Gaps    5;

Qy       3310 ATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTG 3369
              ||||||||||||||||||||||||||||||||:|||||||||:|:||||||:::|||||
Db        427 ATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTG 368

Qy       3370 GCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRG 3429
              | |||| ||| ||  ||    | |||  ||||||||:|||||||| |||:| |||||||:|
Db        367 GTTTTCTAGGACCTAAAGACGGCCCT--AACCCAAGCCCCAGTGTAAAGCTCGCCAACGG 310

Qy       3430 GGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTAC- 3488
              |||||||  |||:||:|||::|||||| |  |||||||:||:|||||:||| |||||||
Db        309 GGCAAGA-GTTTCTTTATATGTCACAG-ATGAAACAGGAATAGCTCTAGGAATCCTTACG 252

Qy       3489 ACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT 3548
              |   :|||:|||||: |||||||||||::||||::|||||:||| ||||:|||||||||
Db        251 AGCAGTCCAAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAAGAAATTGATGTAGT 192

Qy       3549 GGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGA 3608
              ||||||||||:|:|||||||||:||||||||||||||||||||::|||||||||||
Db        191 GGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGA 132

Qy       3609 AGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRG 3668
              ||||||||||||||||  |||:|||||||||||||||||||| |    |:| |||::|
Db        131 AGCAGTTAAAATAATAATGGGAAGAGATCTTACTGTGTGGACATAT----ATGCGAACGG 76

Qy       3669 CATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGC 3728
              |||||| ||||||||||||||||||||||||||||||:|||||||| ||:||||||
Db        75 CATACTTACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAGATATCAGGC 16

Qy       3729 TCTATTACTTGAARG 3743
              ||||||||| ||:|
Db        15 TCTATTACTTAAAGG 1


RESULT 282
AG195659
LOCUS       AG195659                399 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-074N21.T7, genomic survey
            sequence.
ACCESSION   AG195659
VERSION     AG195659.1  GI:45227835
KEYWORDS    GSS.
```

```
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 399)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                Sequencing: T7
             LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-074N21.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"

  Query Match             4.2%;  Score 311.4;  DB 1;  Length 399;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  346;  Conservative   13;  Mismatches   35;  Indels    4;  Gaps    2;

Qy       4200 RATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATC 4259
              :| |||||:||||:||||||||||||||:||||||| :| ||||||||||||:|||||||||
Db             2 AAATATTATTGGCTGTACAGAAACCTAAAGAGGTGACAATCTTACACTGCCGGGGTCATC 61

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
              | |||| ||||:||||||||| |:|||:|||:  |||||| :||||||||||:|||||||
Db            62 ATAAAGAAAAGGAAAGGGAAACAGAAGGGAACCACCAAGCGGATATTGAAGCCAAAAGAG 121

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
              | |||||||||||||||||||| |||||||||| ||   ||| |||||| ||||||||
Db          122 CCGCAAGGCAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATC 181

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
              || ||| |||||||||||||||||||||| || ||||| ||||||||||||||||||||
Db          182 CCCTCCAGGAAACCAAGCCCCAGTACTCAGAAGAAGAAATAGAATGGGGAACCTCACGAG 241

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
              |    ||||| |||||||| |||||| ||||||| |||||| | ||||||||||||||||
Db          242 GACATAGTTTCCTCCCCTCAGGACGGCTAGCCACCGAAGAAAGAAAAATACTTTTGCCTG 301

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACC-TTTCACTTAGGCATCGAT 4555
              || ||| |||||||||||||||||||||||||| |||||| |||||||||||||| |||
Db          302 CAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAAACCTTTTCACTTAGGCATTGAT 361

Qy       4556 AGCACCCATCARATGGCCAAATCATTATTTACTGGACC 4593
              ||||||||||:|||||||||| ||||||||||||||
Db          362 AGCACCCATCAGATGGCCAAATTATTATTTACTGGACC 399


RESULT 283
BI053385/c
```

```
LOCUS       BI053385                334 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-250101-011-D09 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI053385
VERSION     BI053385.1  GI:14460915
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 334)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            250101-011-D09&t3=2001-01-25&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .334
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.1%;  Score 310.1;  DB 1;  Length 334;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  323;  Conservative    2;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       6039 ACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCAC 6098
              |||||||||| |||||||||||||||||||||||||||||||||||||| |   ||||||
Db           333 ACTACATGAACCCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATCCACNCCTCAC 274

Qy       6099 TGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAA 6158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           273 TGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAA 214

Qy       6159 CTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAAT 6218
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           213 CTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAAT 154

Qy       6219 AAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTC 6278
              |||| ||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db           153 AAACCCCACTTCCGTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTC 94

Qy       6279 AAACCTCACCTGTGTAAAATTTAGCAATACT-ACATACACAACCAACTCCCAATGCATCA 6337
              |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||
Db            93 AAACCTCACCTGTGTAAAATTTAGCAATCCTCACATACACAACCAACTCCCAATGCATCA 34
```

```
Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCC 6370
              ||||||||||||||||||||||||||||||||||
Db         33 GGTGGGTAACTCCTCCCACACAAATAGTCTGCC 1
```

RESULT 284
BI053220
LOCUS BI053220 316 bp mRNA linear EST 15-JUN-2001
DEFINITION PM0-GN0210-230101-003-a10 GN0210 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI053220
VERSION BI053220.1 GI:14460750
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 316)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-GN0210-230101-003-a10&t3=2001-01-23&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 316.
FEATURES Location/Qualifiers
source 1. .316
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0210"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 4.1%; Score 309; DB 1; Length 316;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 306; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy       6766 ACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACT 6825
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AACTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACT 61

Qy       6826 GAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACT 6885
              ||||||| |:|||||||||:|||||||||||||||||:||||||:|||||||||||||||
Db         62 GAGAAAGCTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACT 121

Qy       6886 GGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 6945
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        122 GGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 181
```

```
Qy       6946 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTT 7005
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        182 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTT 241

Qy       7006 GTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACT 7065
              |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        242 GTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACT 301

Qy       7066 AAGATCTACCGCAGA 7080
              |||||||||||||||
Db        302 AAGATCTACCGCAGA 316
```

RESULT 285
AQ044697/c
LOCUS AQ044697 398 bp DNA linear GSS 14-JUL-1998
DEFINITION CIT-HSP-2320K8.TR CIT-HSP Homo sapiens genomic clone 2320K8, genomic survey sequence.
ACCESSION AQ044697
VERSION AQ044697.1 GI:3313624
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 398)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2320K8.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .398
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2320K8"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"

Query Match 4.1%; Score 308.9; DB 1; Length 398;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 335; Conservative 32; Mismatches 23; Indels 3; Gaps 3;

```
Qy       3023 AGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTA 3082
              | ||||||||||||||||||||||||||||:|| ||:|| |||||| |:|| ||||||||
Db        390 AAAAACCTTGTGCCATCAAGCCACCCAAGCGCTTTTAAACTTCCTCACCACTTGTGGCTA 331

Qy       3083 CAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAAR 3142
              ||:|||||||||||:|:|:|||||:|||||| |||||||||||||||||||||:||||:
Db        330 CAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAA 271

Qy       3143 ATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCC 3202
              |||||||||||||||||:|||||||||| || ||:::|||||||||||||||||||||||
```

```
Db        270 ATTATCCAAAGGCACCAGGGCCCTCAGTAAGAAATGTATCCAGCCTATACTGGCTTATCC 211

Qy       3203 TCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAW 3262
             |||||:|||||||||||||||| ||| :| :|||||||||:|||:|||:||||||| ||:
Db        210 TCATCCCAAAACCCTAAAGCAAATAAAAG-GTTCCTTGGCATAACAGGCTTCTGCCAAAT 152

Qy       3263 ATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTC 3322
             ||||||| |||||||: |||:||||||||||:|||||:|||||:|||||||||| |||
Db        151 ATGGATT-CCCAGGTACGGCAAAATAGCCAGGCCATTATATACACTAATTAAGGAAGCTC 93

Qy       3323 AGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCC 3382
             |||||||||||||||||||:||| |||||: :|| |||:::|||  ||||||||||||
Db         92 AGAAAGCCAATACCCATTTAGTAAAATGGACTCCTAAAGCAAAAGCAGCTTTCCAGGCCC 33

Qy       3383 CTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTT 3415
                ||| |||||||| |||||||| | |||||||
Db          32 TAAAAAAAGGCCTT-AACCCAAGACTCAGTGTT 1


RESULT 286
B40768/c
LOCUS       B40768                437 bp    DNA     linear   GSS 18-OCT-1997
DEFINITION  HS-1052-B1-A03-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 774 Col=5 Row=B, genomic survey
            sequence.
ACCESSION   B40768
VERSION     B40768.1  GI:2545020
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 437)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 774  row: B  column: 5
            Class: BAC ends
            High quality sequence stop: 437.
FEATURES             Location/Qualifiers
     source          1. .437
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 774 Col=5 Row=B"
                     /sex="M"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.1%;  Score 308.8;  DB 1;  Length 437;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  333;  Conservative   33;  Mismatches   22;  Indels    4;  Gaps    3;

Qy       3032 GTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTC 3091
             ||| |||||||||||||| | :|||||:|||||||| ||:|||||||||||||:||| ||
Db        437 GTGTCATCAAGCCACCCAGGTGCTCTTAAATTTCCTTGCTACCTGTGGCTACAAGGTNTC 378

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
             |||||:|:|:|||||:||||||||||||||||||||||||||||:||||:|||||||||
Db        377 CAAACCAAAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAA 318

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
             ||||||||:| |||||||||||||||:::|||||| ||||| |||||||||| |||:|||
Db        317 AGGCACCAGGACCCTCAGTGAGGAATGTATCCAGGCTATATTGGCTTATCCTTATCCCAA 258

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
             |||||| |||||||||||:| :|||||||||:|||:|||   ||||||||:||||||| |
Db        257 AACCCTGAAGCAACTAAGAGCGTTCCTTGGCATAACAGG-AACTGCCGAATATGGATT-C 200

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
             ||||||: |||:|||||||||| :|||||:|||||:|||||||||||||||||||||||
Db        199 CCAGGTACGGCGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCA 140

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
             |||||||| ||:|||||||||:|:||||||:::|||||||| |||||||| ||  |||
Db        139 ATACCCATCTAGTAAGATGGACACCTGAAGCAGAAGTGGCTNTCCAGGCCCTAAAGAAAG 80

Qy       3392 GCCTTAAACCCAAGYCCCAGTGTTAAGYTTGC 3423
                |||  ||||||||:|||||||||||||:||||
Db          79 CCCT--AACCCAAGCCCCAGTGTTAAGCTTGC 50


RESULT 287
AL598182/c
LOCUS       AL598182                408 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp313H1319_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313H1319 5', mRNA sequence.
ACCESSION   AL598182
VERSION     AL598182.1  GI:15160873
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp313H1319) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313H1319"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="313 (synonym: hlcc2)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"

  Query Match             4.1%;  Score 308.1;  DB 1;  Length 408;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches  373;  Conservative    7;  Mismatches   26;  Indels    7;  Gaps    5;

Qy        821 TTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGA 880
              |||||||||||||||||| ||| ||||||||||||||||| |||||||||||||||
Db        408 TTCTTTTCATTAAGAGACACCTCGCAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGA 349

Qy        881 AGCCTTCAGAGTCTACCTCCCTATCCCAGCAT--CCCCGACTCCTTCCCCAMYTAATAAG 938
              |||| |||||||||||||||||||||||  |  ||||||||||||||::||| |||
```

```
Db        348 AGCCCTCAGAGTCTACCTCCCTATCCCAGTGTCCCCCCGACTCCTTCCCCAACTAACAAG 289

Qy        939 GACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCA 998
              |||||||||| |||||||| ||||| |||||||| ||||||| ||||||||| ||||||
Db        288 GACCCCCCTTTAACCCAAACGGTCC-AAAGGAGACAGACAAAGGGGTAAACAATGAACC- 231

Qy        999 AAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGG 1058
              |||||||||||||||||| |||||| ||||| |||||||||||||||| ||||||||||
Db        230 -AGAGTGCCAATATTCCCCGATTATGCCCCCT-CCAAGCAGTGGGAGGAGGAGAATTCGG 173

Qy       1059 CCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTT 1118
              |||||||||||||||||| |:||||::| || |||||||||||||||| |||| |||  |
Db        172 CCCAGCCAGAGTGCATGTACCTTTTTCTGTCTCAGACTTAAAGCAAATTAAAATAGATCT 113

Qy       1119 AGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATT 1178
              ||||||||||||||||||:|||||||||||||||: |||||||||||||||||||||||
Db        112 AGGTAAATTCTCAGATAACCCTGATGGCTATATTGACGTTTTACAAGGGTTAGGACAATC 53

Qy       1179 CTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACC 1231
              |||||||||||||||||||||||| |||| ||||||| ||||||||||||||
Db         52 CTTTGATCTGACATGGAGAGATATA-ATGTTACTGCTAGATCAGACACTAACC 1
```

RESULT 288
AA368311
LOCUS AA368311 364 bp mRNA linear EST 21-APR-1997
DEFINITION EST79542 Placenta I Homo sapiens cDNA similar to similar to endogenous retrovirus ERV9, mRNA sequence.
ACCESSION AA368311
VERSION AA368311.1 GI:2020630
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 364)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
PUBMED 7566098
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13-21.
FEATURES Location/Qualifiers
source 1. .364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):172998"
/db_xref="taxon:9606"
/tissue_type="placenta"
/dev_stage="fetus"
/clone_lib="Placenta I"
/note="Organ: placenta; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: EcoRI"

Query Match 4.1%; Score 304.1; DB 1; Length 364;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 349; Conservative 1; Mismatches 7; Indels 5; Gaps 4;

```
Qy       2238 TTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 2297
              ||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          6 TTCTCCCAGCCACTAAGTTATG-ACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 64

Qy       2298 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 124

Qy       2358 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAA 2417
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        125 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAA 183

Qy       2418 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCT 2477
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        184 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCCGTCCT 242

Qy       2478 GTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGA 2537
              |||||||||||||||||||||||| |||||||||||||||||||||||||| ||||||
Db        243 GTTCAAGTTAAACTAAAGGATTCCACCTCCTTTCCCTACCAAAGGCAGTACCCNCTCAGA 302

Qy       2538 CCCAAGGCCCAACAA--GGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
              |||||||||||||||  || |||||||||||||||||||| | |||| ||||||| |||
Db        303 CCCAAGGCCCAACAAAGGGCTTCCAAAAGATTGTTAAGGACCTTAAAGNCCAAGGCCTAG 362

Qy       2596 TA 2597
              ||
Db         363 TA 364
```

RESULT 289
AQ549534/c
LOCUS AQ549534 366 bp DNA linear GSS 28-MAY-1999
DEFINITION RPCI-11-423L5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-423L5, genomic survey sequence.
ACCESSION AQ549534
VERSION AQ549534.1 GI:4908711
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 366)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

```
          Seq primer: T7
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7662316"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-423L5"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             4.1%;  Score 303.7;  DB 1;  Length 366;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  319;  Conservative   19;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3693 GTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCT 3752
             |||||||||||||||||:|||||||||||:||||||||||||||||||:||||||||||
Db        366 GTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCT 307

Qy       3753 GCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAA 3812
             | :||||   |||||||||||||||| ||||:  | |||||||||| |||| ||||| |
Db        306 GTGACTGCATACTTGTGCAACTCTTAATCCAGCAACGTTTCTTCCAGTCAATAAAGAATA 247

Qy       3813 GATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYT 3872
             ||||:||:||  |||||||| |:||||| ||||||| || ||  ||||||||||||||:|
Db        246 GATAGAACATTTCTGTCAACGAGTAATTGCTCAAACATACGCTGCTCGAGGGGACCTTTT 187

Qy       3873 AGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAG 3932
             |||:|||||:||||||||||: ||| ||| ||||||||||||||||||||||||||||||
Db        186 AGAGGTTCCCTTGACTGATCCCAACC-TCAGCTTGTATACTGATGGAAGTTCCTTTGTAG 128

Qy       3933 AAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTA 3992
             |||| |||||| ||||||:| ||||||||||||||||||||||||||:|| ||||||
Db        127 AAAAGGGACTTTGAAAAGCGAGGTATGCAGTGGTCAGTGATAATGGAATACTTAAAAGTA 68

Qy       3993 ATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCAC 4052
             ||||||||||||||||||||||||:|| |:|||||||||||||:|| :|:||||||
Db          67 ATCCCCTCACTCCAGGAACTAGTGCTCAGTTGGCAGAACTAATAGCCCTCTCTTGGGCAC 8

Qy       4053 TAGAATT 4059
             || ||||
Db           7 TAAAATT 1


RESULT 290
BI056596/c
LOCUS       BI056596                313 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-210201-013-H02 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056596
VERSION     BI056596.1  GI:14464126
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            210201-013-H02&t3=2001-02-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 312.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.0%;  Score 301.6;  DB 1;  Length 313;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches  301;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6067 CGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 6126
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CCCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 245

Qy       6127 CCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTA 6186
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        244 CCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTA 185

Qy       6187 CCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT 6246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 CCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT 125

Qy       6247 CTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAAT 6306
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAAT 65

Qy       6307 ACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTC 6366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          64 ACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTC 5

Qy       6367 TGCC 6370
             ||||
Db           4 TGCC 1


RESULT 291
AA663366
LOCUS       AA663366                342 bp    mRNA    linear   EST 12-NOV-1997
DEFINITION  ab82d03.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
            IMAGE:853445 3' similar to TR:G300482 G300482 POL=REVERSE
            TRANSCRIPTASE HOMOLOG {RETROVIRAL ELEMENT} ;, mRNA sequence.
ACCESSION   AA663366
VERSION     AA663366.1  GI:2617357
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
```

```
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:853445"
                     /sex="mixed"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene fetal retina 937202"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Pooled
                     retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

  Query Match             4.0%;  Score 300.8;  DB 1;  Length 342;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  296;  Conservative   26;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       3991 TAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGC 4050
             ||||||||||||||||||||||||||||||:||||:|||| ||||||||||:|||:|  |||
Db          1 TAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGGACTAATAGCCCTCACTCAGGC 60

Qy       4051 ACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCT 4110
             |||||||||||||||:| ||||||:|||||||||||||||||:|:||||||:||||
Db         61 GCTAGAATTAGGAGAAGGAGAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCT 120

Qy       4111 AGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAAC 4170
             |||| |||| ||||| |:: |||||||:|:||||||||||||||||||: ||:||||
Db        121 AGTCCTCCAAGCCCAGGCAACAATATGGAGAGAAAGGGAATTCCTAACTTCCAAGGGAAC 180

Qy       4171 ACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGA 4230
             | ||||||:|||||||||||||||||:|||||||:|||  |||||||| ||||:|||
Db        181 ATCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTGGTACAGAACCCTAAAGA 240

Qy       4231 GGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAA 4290
             ||||| ||||||||||||:|||||||||| |||||||||:||||||||||:|||:|||
Db        241 GGTGGTAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAA 300

Qy       4291 YTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             :|||||||| :||||| ||||:|||||||| ||||| |||||
Db        301 CTGCCAAGCGGATATTAAAGCAAAAAGAGCCGCAAGACAGGA 342


RESULT 292
R27389/c
LOCUS       R27389                 404 bp    mRNA    linear   EST 24-APR-1995
DEFINITION  yh46a09.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:132760 3', mRNA sequence.
ACCESSION   R27389
VERSION     R27389.1  GI:783524
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 793
            High quality sequence stops: 250
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 793   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 250.
FEATURES             Location/Qualifiers
     source          1. .404
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:538399"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:132760"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             4.0%;  Score 300.8;  DB 1;  Length 404;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  333;  Conservative    8;  Mismatches   31;  Indels    4;  Gaps    2;

Qy       7210 CCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGA--TGGGGGACTGAGAGACAGGA 7267
             ||  || |  || || |    | ||| |||||||| |   ||||  ||| ||||||||
Db        376 CCTCCCCCAACAGCACCTTAGGGTTTTCCCTGTTGAAAATGGGGGACTTGANAGACAGGA 317

Qy       7268 CTAGCTGGATTTCCTAGGCTGAYTAAGAAT--CCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||| |||| ||||||||||:  |  |   ||:||||||||:|||||||||||||||
Db        316 CTAGCTNGATTCCCTAGGCTGACCTAAGAATCCCCTAAGCCTAGCTGGGAAGGTGACCAC 257

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        256 ATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCANACCTGACCAATCAGAGAGCTCAC 197

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             | |||||||||||||||| |||||||||||||||||||||||:||||||:||||
Db        196 TNAAATGCTAATTAGGCAAANACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 137

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        136 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCA 77

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 ANCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 17

Qy       7566 TAAATCTTGCARCTGC 7581
             ||||||||||:||||
Db         16 TAAATCTTGCAACTGC 1
```

```
RESULT 293
AQ192721/c
LOCUS       AQ192721                383 bp    DNA     linear   GSS 04-NOV-1998
DEFINITION  HS_2251_B2_F12_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2251 Col=24 Row=L, genomic survey
            sequence.
ACCESSION   AQ192721
VERSION     AQ192721.1  GI:3591343
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2251  row: L  column: 24
            Class: BAC ends
            High quality sequence stop: 383.
FEATURES             Location/Qualifiers
     source          1. .383
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2251 Col=24 Row=L"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             4.0%;  Score 300.5;  DB 1;  Length 383;
  Best Local Similarity   84.3%;  Pred. No. 0;
  Matches  328;  Conservative   32;  Mismatches   22;  Indels    7;  Gaps    3;

Qy       3029 CTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGT 3088
             || |||||||| |||||||||||:|||||:|||||||| ||:|||||||||||||:|||
Db        383 CTGGTGCCATCGAGCCACCCAAGCACTCTTAAATTTCCTTGCTACCTGTGGCTACAAGGT 324

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATC 3148
             |||||||| |:| |||||:| |||||||||||||||||| |||||||:||||:||||||
Db        323 TTCCAAACAAAATGCTCAGCACTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATC 264

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208
             ||||||||| |:|||||||||||||||:::|||||||||||||||||||||||||||:
Db        263 CAAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCC 204

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
             ||||||||||||||||||||:|::|||||||||:|||:| |:|||||||||:||||||
Db        203 CAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGAT 144

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
             | |||||||:  ||:|||||| ||||:|||||:| |||:    ||||||||||||||||
Db        143 T-CCCAGGTACAGCAAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCAGAAAG 89

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
             |||||||||||||:||||||||:|:||||||:::|||||||||| ||||||   ||
Db         88 CCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCATAAAGA 29

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             | | |||  |||||||||:||||||| |||
Db         28 AGGCCCT--AACCCAAGCCCCAGTGGTAA 2


RESULT 294
AA781468
LOCUS       AA781468                417 bp    mRNA    linear   EST 31-DEC-1998
DEFINITION  aj26h03.s1 Soares_testis_NHT Homo sapiens cDNA clone 1391477 3',
            mRNA sequence.
ACCESSION   AA781468
VERSION     AA781468.1  GI:2840799
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1116   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 315.
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="1391477"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares_testis_NHT"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from mRNA obtained from
                     Clontech Laboratories, Inc., and primed with  a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. .
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "

  Query Match             4.0%;  Score 300.5;  DB 1;  Length 417;
  Best Local Similarity   84.3%;  Pred. No. 0;
  Matches  355;  Conservative    2;  Mismatches   55;  Indels    9;  Gaps    2;

Qy         17 AACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGTATGGG 76
              || |||||||||||||||||||| |||  |||| ||||||| || ||||||||||||||
Db          6 AAACCAGGCATTCGAGCTGGCAACGGCA-ACCCCATTTGGGTTCCCTCCCTTTGTATGGG 64

Qy         77 AGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGT 136
              |||| | |         |||||||||||||||||| ||||||||||||||||||  |||
Db         65 AGCTCTGT--------TTTCACTCTATTAAATCTTGGAACTGCACTCTTCTGGTCTGTGT 116

Qy        137 TTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGC 196
              || ||| |||| |||||| |||||||||| |||||||||||||||| | ||| |    ||
Db        117 TTGTTATGGCTGGAGCTGGGCTTTTGCTCGCCGTCCACCACTGCTGCTGGCCGCTGTTGC 176

Qy        197 ANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATC 256
              | ||| |||||||||| |||||||||  ||||| |||||||||| ||||| |||||||||
```

```
Db        177 AGACCCGCCGCTGACTTCCATCCCTCCAGATCCAGCAGGGTGTCTGCTGTACTCCTGATC 236

Qy        257 CAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGG 316
              || ||:|| ||||||||| ||| | || ||||||||||||||    ||  ||||||| |
Db        237 AAGTGAGGCACCCATTGCCCCTCTCCATCGGGCTAAAGGCTTGCATGTGCTCCTGCACAG 296

Qy        317 CTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTC 376
              |||||||||||||| || |||||||||| ||||||| | ||||||||||||||||||||
Db        297 TTAAGTGCCTGGGTTCGTCCTAATTGAGCTAAACACTAATTACTGGGTTCCATGGTTCTC 356

Qy        377 TTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCAT 436
              |||  |||||||| ||||||||||:| |||||||||| ||||||||||||||||||||
Db        357 TTCCATGACCCACTGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGATTCCAT 416

Qy        437 T 437
              |
Db        417 T 417


RESULT 295
H12439
LOCUS       H12439                  414 bp    mRNA    linear   EST 27-JUN-1995
DEFINITION  yj11h10.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:148483 5', mRNA sequence.
ACCESSION   H12439
VERSION     H12439.1  GI:877259
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 678
            High quality sequence stops: 257
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 678   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 257.
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:560252"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:148483"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo. "

  Query Match             4.0%;  Score 300.2;  DB 1;  Length 414;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  369;  Conservative    5;  Mismatches   24;  Indels   16;  Gaps    5;

Qy       5053 CAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAAC 5112
              |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db          1 CAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAAC 60

Qy       5113 TTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACC 5172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACC 120

Qy       5173 AATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCA 5232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCA 180

Qy       5233 AATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAAC 5292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAAC 240

Qy       5293 TATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTA- 5351
              ||||||||::|||||||||||||||||||||||:||||||||||||||||||||||||
Db        241 TATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAG 300

Qy       5352 GATACATCCTGGG----AAGGACCCTACCCAGT---CATTTTATYTACCCCAACTGCGG- 5403
              |||||||||||||      |  || ||||||||     |||  |:||||||| |||||
Db        301 GATACATCCTGGGGGAAGGGACCCTTACCCAGTTCATTTTTATTTTACCCCAATTGCGGG 360

Qy       5404 -TTAAAGTGGCTGGAGTGGAG------TCTTGGATACATCACACTTGAGTCAAA 5450
              |||||||||   | | |  |       |   || |||||||| ||||||  ||
Db        361 TTTAAAGTGGGCTGGGAGTTGGAGTTTTTNGGGTTACATCACATTTGAGTTCAA 414


RESULT 296
AQ412083
LOCUS       AQ412083                398 bp    DNA     linear   GSS 23-MAR-1999
DEFINITION  RPCI-11-195J9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-195J9,
            genomic survey sequence.
ACCESSION   AQ412083
VERSION     AQ412083.1  GI:4471571
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 398)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="GDB:7574720"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-195J9"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             4.0%;  Score 299.9;  DB 1;  Length 398;
  Best Local Similarity   83.2%;  Pred. No. 0;
  Matches  331;  Conservative   15;  Mismatches   47;  Indels    5;  Gaps    1;

Qy       4149 AATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAY 4208
             ||||||||||||:|||:||||| ||||||:||||||||||||||||||:| |||||:
Db          1 AATTCCTAACTTCCGAGGGAACATCTATCAAACATCAGGAAGCCATTAGGAGACTATTAT 60

Qy       4209 TGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAA 4268
             ||||:||||||||||||:||||||||:|||||||||| |: |||||||| |||||||
Db         61 TGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTTCCAGGGTCATCAGAAAGGAA 120

Qy       4269 AGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGC 4328
             ||:||||||||||:|||:|||:  |||||||:||||||||| :|||||| ||| |||||
Db         121 AGGAAAGGGAAATAGAAGGGAACCACCAAGCAGATATTGAAGACAAAAGAACTGTAAGGC 180

Qy       4329 AGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGG 4388
               | ||||| |||||||||||||||||  ||   ||| |  ||| | |||||||| | ||||
Db         181 GGTACCCTTCATTAGAAATGCTTATAGAATGACCCCTCATATGGAGTAATCCCATACGGG 240

Qy       4389 AAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC-----GAGGCAG 4443
             ||||||||||||||||||||||  |||||||||||||||| |||||     ||   ||
Db         241 AAACCAAGCCCCAGTACTCAGCAAAAGAAACAGAATGGGGAAACTCACTAGAGGACATAG 300

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             |||  |||||||| ||| || || ||   | |||||||  |||||||| || |||||  |||
Db         301 TTTCCTCCCCTCAGGATGGCTAACCCATTAAGAAGGAAAAATACTTCTGACTGCAGCTAA 360

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTC 4541
                |||| |||||||||||||| |||||| ||||||||||
Db         361 CCAACGGAAATTACTTAAAAACCTTCACCAAACCTTTC 398


RESULT 297
BF990902
LOCUS       BF990902                322 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  RC6-GN0070-271000-023-H12 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF990902
VERSION     BF990902.1  GI:12397227
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 322)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            271000-023-H12&t3=2000-10-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 22
            High quality sequence stop: 275.
FEATURES             Location/Qualifiers
     source          1. .322
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.0%;  Score 298.8;  DB 1;  Length 322;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches  294;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 80

Qy       6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754
             ||||||||||||||||||||||||||:||||||||||||||||||:||||||||||
Db         81 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 140

Qy       6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814
             |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 200

Qy       6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874
             ||||||||||||||||||||:|||||||||:||||||||||||||||:||||||:||||
Db        201 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 260

Qy       6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        261 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 320

Qy       6935 TA 6936
              ||
Db        321 TA 322


RESULT 298
BQ365206/c
LOCUS       BQ365206                310 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC6-GN0070-220800-021-D11 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365206
VERSION     BQ365206.1  GI:21040718
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
```

```
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            220800-021-D11&t3=2000-08-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 19
            High quality sequence stop: 310.
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.0%;  Score 297.6;  DB 1;  Length 310;
  Best Local Similarity   95.8%;  Pred. No. 0;
  Matches  297;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        310 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAGGAACTAAA 251

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 191

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGG 6764
             |||||||||||||| ||:||||||||||||||||||:||||||||||||:||||||||
Db        190 AGCAGCAGTAGTCCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGG 131

Qy       6765 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 6824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 71

Qy       6825 TGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACAC 6884
             ||||||||||:|||||||||:||||||||||||||||:||||||:||||||||||||
Db          70 TGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACAG 11

Qy       6885 TGGACCCTGG 6894
             ||| ||  ||
Db          10 TGGCCCTAGG 1


RESULT 299
BF992540/c
LOCUS       BF992540                 298 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-301000-203-g12 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF992540
VERSION     BF992540.1  GI:12398863
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 298)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            301000-203-g12&t3=2000-10-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 259.
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             4.0%;  Score 297.2;  DB 1;  Length 298;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  296;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4887 AGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCA 4946
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 AGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCA 239

Qy       4947 GGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGT 5006
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGT 179

Qy       5007 CGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCT 5066
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        178 CGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCT 119

Qy       5067 CACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGC 5126
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGC 59

Qy       5127 AGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 5184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          58 AGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 1


RESULT 300
R27388
LOCUS       R27388                   407 bp    mRNA    linear   EST 24-APR-1995
DEFINITION  yh46a09.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:132760 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
            POLYPROTEIN - ;, mRNA sequence.
ACCESSION   R27388
VERSION     R27388.1  GI:783523
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 407)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 793
            High quality sequence stops: 305
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 793   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 305.
FEATURES             Location/Qualifiers
     source          1. .407
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:538399"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:132760"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             3.9%;  Score 295.7;  DB 1;  Length 407;
  Best Local Similarity   89.4%;  Pred. No. 0;
  Matches  320;  Conservative    6;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       6867 AGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCT 6926
             || :|||||||||||||||||||||||||||||||| ||||:||||||||||||||||||
Db         13 AGGAGAGCTTCGAAACACTGGACCCTGGGGCCTCCTNAGCCAATGGATGCCCTGGATTCT 72

Qy       6927 CCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTT 6986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 CCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTT 132

Qy       6987 TRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCC 7046
             |:|||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        133 TAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCC 192

Qy       7047 CAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCT-GGACCGGCCTGYTAGCCCAC 7105
             |||||||||||||||||||||||||||||||||||||| ||||||||||:||||||||
Db        193 CAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGGACCGGCCTGCTAGCCCAC 252

Qy       7106 GATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTC 7165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTC 312

Qy       7166 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAAC 7223
             ||||||| || | |         | ||  ||   || :|  || |   |  |||||
Db        313 TACTACGGNCCCAATTCAGGCAGGGAAGGCAGTTAGGAGGCGGGTCGTTCGGGCCAAC 370


RESULT 301
AQ020123
LOCUS       AQ020123                358 bp    DNA     linear   GSS 09-JUN-1998
DEFINITION  CIT-HSP-2303O3.TR CIT-HSP Homo sapiens genomic clone 2303O3,
            genomic survey sequence.
ACCESSION   AQ020123
VERSION     AQ020123.1  GI:3198859
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2303O3.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .358
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2303O3"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             3.9%;  Score 294.7;  DB 1;  Length 358;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  325;  Conservative    0;  Mismatches   33;  Indels    1;  Gaps    1;

Qy       1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
             ||||||||||||||||||||| ||||  |||||||||||| ||| ||||||||||||||
Db          1 AATAGAGATCAGGAGGAGCAGGTGGAATGGGACAAACGGGA-TAAGAAAAAGGCCACCGC 59

Qy       1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
             |||||||||| ||||||||||||  ||||||||||||||||||| ||||||| |||||||
Db         60 TTTAGTCATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGC 119

Qy       1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
             ||||  |||||||||||||||||||||||||| ||||||||||||||| |||||||||
Db        120 AAATCAAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTGTAAAAAAGA 179

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||| | ||||| |||||| |||| |||||||||||||||| |||||||| ||||||
Db        180 TTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGTCAAGGGAGTCACTG 239

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             |||||||||||||||||||||  |||||| | ||||||||||||||||||| ||||
Db        240 GAAGGCCCACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTAACCACATGAG 299
```

```
Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
Db        300 CCGGTAGCAGGACTGAGGGTGCCCGGGGCAAGCACCAGCCCATACCATCACCGTCACAG 358
```

RESULT 302
AA729556/c
LOCUS AA729556 328 bp mRNA linear EST 22-JAN-1998
DEFINITION nx58c05.s1 NCI_CGAP_Alv1 Homo sapiens cDNA clone IMAGE:1266440, mRNA sequence.
ACCESSION AA729556
VERSION AA729556.1 GI:2750915
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 433 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .328
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1266440"
/tissue_type="alveolar rhabdomyosarcoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Alv1"
/note="Vector: pAMP10; mRNA made from alveolar rhabdomyosarcoma, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

Query Match 3.9%; Score 293.6; DB 1; Length 328;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 301; Conservative 9; Mismatches 3; Indels 2; Gaps 1;

```
Qy       7268 CTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACAT 7327
Db        321 CTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACAT 262

Qy       7328 CCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTA 7387
Db        261 CCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTA 202

Qy       7388 AAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAG 7447
Db        201 AAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAG 142

Qy       7448 CACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAAC 7507
Db        141 CACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGGACCGGCAACGGCAAC 82

Qy       7508 CCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTA 7567
Db         81 CCC--TTTGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTA 24

Qy       7568 AATCTTGCARCTGCR 7582
Db         23 AATCTTGCAACTGCA 9
```

RESULT 303
AQ033092
LOCUS AQ033092 377 bp DNA linear GSS 02-JUL-1998
DEFINITION HS_2227_A1_B07_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2227 Col=13 Row=C, genomic survey sequence.
ACCESSION AQ033092
VERSION AQ033092.1 GI:3285280
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 377)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2227 row: C column: 13
Class: BAC ends
High quality sequence stop: 377.
FEATURES Location/Qualifiers
source 1. .377
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2227 Col=13 Row=C"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

Query Match 3.9%; Score 293.5; DB 1; Length 377;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 322; Conservative 30; Mismatches 22; Indels 9; Gaps 3;

```
Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db          1 AAGACTTTTCTTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAA 59

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db         60 GTCCGAGGGACCAGCTTGCAACCCGTGGCATACATGAGTAAGGAAATTGATGTAGTGGCA 119

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        120 AAGGGCTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCA 179

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        180 GTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATCATGTGAATGGCATA 239

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGC---T 3729
Db        240 CTCACTGCTAAAGGGGACTTCNGGTTTGCAGACAACCATCTGCTTAAATATCAAGCCAGT 299
```

```
Qy       3730 CTATTACTTGAAGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              ||||||||||||:||||||||||||:||||  ||||||||||||||||     ||:| ||
Db        300 CTATTACTTGAAAGGCCAGTGCTGCAACTGCACACTTGTGCAACTCTT-----AGCCACA 354

Qy       3790 TTTCTTCCAGACAATGAAGAAAA 3812
              |||||||||||||||| ||||
Db        355 TTTCTTCCAGACAATGAATAAAA 377


RESULT 304
T93661
LOCUS       T93661                   374 bp    mRNA    linear   EST 23-MAR-1995
DEFINITION  ye05g09.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:116896 5', mRNA sequence.
ACCESSION   T93661
VERSION     T93661.1  GI:726834
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 834
            High quality sequence stops: 308 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 834   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 308.
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:472513"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:116896"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             3.9%;  Score 290.8;  DB 1;  Length 374;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  327;  Conservative    2;  Mismatches   12;  Indels    4;  Gaps    3;

Qy       6085 AATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATA 6144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATA 60

Qy       6145 TGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAAC 6204
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 TGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAAC 120

Qy       6205 TTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAA 6264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 TTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAA 180

Qy       6265 ATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAAC 6324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAAC 240

Qy       6325 TCCCAATGCATCAGGT-GGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA-GGAA 6382
              | |||||||||||||| |||||||| |||||||||||||||||||||||||| | ||||
Db        241 TTCCAATGCATCAGGTGGGGTAACTNCTCCCACACAAATAGTCTGCCTACCCTTAGGGAA 300

Qy       6383 TATTTTTTGTCTGTGG--TACCTCAGCCTATCGTTGTTTGAATGG 6425
              ||||||||||||||||    |    | |||| |||||||| ||||
Db        301 TATTTTTTGTCTGTGGGTACCTCAGGNCTATTGTTGTTTGNATGG 345


RESULT 305
AQ024733/c
LOCUS       AQ024733                 332 bp    DNA     linear   GSS 23-JUN-1998
DEFINITION  HS_2166_B2_MR_D12 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2166 Col=24 Row=H, genomic survey
            sequence.
ACCESSION   AQ024733
VERSION     AQ024733.1  GI:3243328
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2166  row: H  column: 24
            Class: BAC ends
            High quality sequence stop: 332.
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2166 Col=24 Row=H"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             3.9%;  Score 289;  DB 1;  Length 332;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  295;  Conservative   31;  Mismatches    6;  Indels    2;  Gaps    2;

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||:|||::|||||||||||||||| |||||||||||||:|||||:|||||||||
Db        332 TTGGCCGCCCATTCAGAAACCTTGTGCCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 273
```

```
Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
             ||:|||||||||||||:||||||||||:|:|:|||||:|||||||||||||| |||||
Db        272 GCCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGTTTAAA 213

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
             ||||||||:||||:||||||||||||||||:|||||||||| ||||:::|||||||||
Db        212 TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGGGGAATGTATCCAGCCT 153

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
             ||||| ||||||||||||:| ||||||||||||||||||:|::|||||||||:|||:|
Db        152 ATACT-GCTTATCCTCATCCCCAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACA 94

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
             ||:|||||||||:|||||||| ||||||: |||:|||||||||| :|||||:|||||:|
Db          93 GGTTTCTGCCGAATATGGATT-CCCAGGTACGGCAAAATAGCCAGACCATTATATACACT 35

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTA 3342
             |||||||||||||||||||||||||||||||||
Db          34 AATTAAGGAAACTCAGAAAGCCAATACCCATTTA 1
```

RESULT 306
BQ024998/c

```
LOCUS       BQ024998                347 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-1-BB1p-aty-h-05-0-UI.s1 NCI_CGAP_Pl6 Homo sapiens cDNA clone
            UI-1-BB1p-aty-h-05-0-UI 3', mRNA sequence.
ACCESSION   BQ024998
VERSION     BQ024998.1  GI:19760284
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 347)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            The following repetitive elements were found in this cDNA
            sequence: 18-45, >LTR17#LTR/Retroviral (matched compliment) 37-339,
            >HERV17#LTR/Retroviral (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .347
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BB1p-aty-h-05-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="Full Term"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Pl6"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl6 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human full term
                     TAG_LIB=UI-1-BB1p
                     TAG_SEQ=AGGAA"
```

```
  Query Match             3.8%;  Score 288.1;  DB 1;  Length 347;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  295;  Conservative    7;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5333 CTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTAC 5392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        339 CTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTAC 280

Qy       5393 CCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATC 5452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATC 220

Qy       5453 CTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACC 5512
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        219 CTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGATCC 160

Qy       5513 TCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATA 5572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 TCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATA 100

Qy       5573 AATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCA 5632
             ||| ||||||||:||||||||||||||||||||||:||:||||:|||:|||||:|||||
Db          99 AAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCA 41

Qy       5633 CTCT 5636
             ||||
Db          40 CTCT 37
```

RESULT 307
BI055459/c

```
LOCUS       BI055459                315 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-060201-012-H05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI055459
VERSION     BI055459.1  GI:14462989
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            060201-012-H05&t3=2001-02-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
```

```
          High quality sequence stop: 242.
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 286.4;  DB 1;  Length 315;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  298;  Conservative    2;  Mismatches    5;  Indels    2;  Gaps    1;

Qy       6068 GCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACC 6127
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        315 GCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGNCCAAAACN 256

Qy       6128 CTA--CTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGT 6185
             | |    ||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        255 CAAACACAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGT 196

Qy       6186 ACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACC 6245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 ACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACC 136

Qy       6246 TCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA 6305
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA 76

Qy       6306 TACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGT 6365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 TACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGT 16

Qy       6366 CTGCCTA 6372
             |||||||
Db          15 CTGCCTA 9


RESULT 308
BI056080/c
LOCUS       BI056080                315 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-150201-012-H05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056080
VERSION     BI056080.1  GI:14463610
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            150201-012-H05&t3=2001-02-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 242.
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 286.4;  DB 1;  Length 315;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  298;  Conservative    2;  Mismatches    5;  Indels    2;  Gaps    1;

Qy       6068 GCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACC 6127
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        315 GCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGNCCAAAACN 256

Qy       6128 CTA--CTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGT 6185
             | |    ||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        255 CAAACACAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGT 196

Qy       6186 ACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACC 6245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 ACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACC 136

Qy       6246 TCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA 6305
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA 76

Qy       6306 TACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGT 6365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 TACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGT 16

Qy       6366 CTGCCTA 6372
             |||||||
Db          15 CTGCCTA 9


RESULT 309
BI055458/c
LOCUS       BI055458                331 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-060201-012-E05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI055458
VERSION     BI055458.1  GI:14462988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 331)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
```

```
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
    PUBMED  10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            060201-012-E05&t3=2001-02-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 20
            High quality sequence stop: 256.
FEATURES             Location/Qualifiers
     source          1. .331
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 285.9;  DB 1;  Length 331;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  300;  Conservative    2;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       6064 ACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 ACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 272

Qy       6124 AACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183
             || |  || |||| |||||||||||||||||||||:||||||||||||||| |||
Db        271 AANCAAACAAACTATTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATACCT 212

Qy       6184 GTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGA 6243
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTACCCGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGA 152

Qy       6244 CCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGC 6303
             |||||||||||| ||:||||||||||||||||||||||||||||||||||||||||||||
Db        151 CCTCTTGTTTCCGATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGC 92

Qy       6304 AATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATA 6363
             |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db         91 AATACTACATACAC-ACCAACTCCAAATGCATCAGGTGGGTAACTCCTCCCACACAAATA 33

Qy       6364 GTCTGCCTACCC 6375
             ||||||||||||
Db         32 GTCTGCCTACCC 21


RESULT 310
BI056079/c
LOCUS       BI056079                331 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-150201-012-E05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056079
VERSION     BI056079.1  GI:14463609
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 331)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
    PUBMED  10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            150201-012-E05&t3=2001-02-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 20
            High quality sequence stop: 256.
FEATURES             Location/Qualifiers
     source          1. .331
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 285.9;  DB 1;  Length 331;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  300;  Conservative    2;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       6064 ACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 ACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 272

Qy       6124 AACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183
             || |  || |||| |||||||||||||||||||||:||||||||||||||| |||
Db        271 AANCAAACAAACTATTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATACCT 212

Qy       6184 GTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGA 6243
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTACCCGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGA 152

Qy       6244 CCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGC 6303
             |||||||||||| ||:||||||||||||||||||||||||||||||||||||||||||||
Db        151 CCTCTTGTTTCCGATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGC 92

Qy       6304 AATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATA 6363
             |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db         91 AATACTACATACAC-ACCAACTCCAAATGCATCAGGTGGGTAACTCCTCCCACACAAATA 33

Qy       6364 GTCTGCCTACCC 6375
             ||||||||||||
Db         32 GTCTGCCTACCC 21


RESULT 311
```

```
BI055456/c
LOCUS       BI055456                300 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-060201-012-D05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI055456
VERSION     BI055456.1  GI:14462986
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            060201-012-D05&t3=2001-02-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 232.
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 284.8;  DB 1;  Length 300;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  289;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6071 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTA 6130
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||    |
Db        300 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATAAGGTCTCGGCCCAAAACAAAA 241

Qy       6131 CTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTG 6190
             | ||||||||||||||||||||||||||||||:|||||||||||||| ||||||||||
Db        240 CAAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAAAGCCTGTACCTG 181

Qy       6191 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 6250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 121

Qy       6251 TTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 6310
             |||| |||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTTCAAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 61

Qy       6311 CATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 6370
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||

Db         60 CATACACAACCAACTCCAAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 1


RESULT 312
BI056077/c
LOCUS       BI056077                300 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  RC5-GN0282-150201-012-D05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI056077
VERSION     BI056077.1  GI:14463607
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-
            150201-012-D05&t3=2001-02-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 232.
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0282"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 284.8;  DB 1;  Length 300;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  289;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6071 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTA 6130
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||    |
Db        300 TGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATAAGGTCTCGGCCCAAAACAAAA 241

Qy       6131 CTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTG 6190
             | ||||||||||||||||||||||||||||||:|||||||||||||| ||||||||||
Db        240 CAAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAAAGCCTGTACCTG 181

Qy       6191 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 6250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 AACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTG 121

Qy       6251 TTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 6310
             |||| |||:||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        120 TTTCAAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTA 61

Qy       6311 CATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 6370
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         60 CATACACAACCAACTCCAAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCC 1


RESULT 313
BI053219
LOCUS       BI053219                289 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  PM0-GN0210-230101-003-a05 GN0210 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI053219
VERSION     BI053219.1  GI:14460749
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-GN0210-
            230101-003-a05&t3=2001-01-23&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 289.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0210"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.8%;  Score 283.6;  DB 1;  Length 289;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches  280;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6766 ACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACT 6825
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AACTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACT 61

Qy       6826 GAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACT 6885
              |||||||||:|||||||||:||||||||||||||||:||||||:|||||||||||||||
Db         62 GAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACT 121

Qy       6886 GGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 6945
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||

Db        122 GGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTA 181

Qy       6946 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTT 7005
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        182 GCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTT 241

Qy       7006 GTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATG 7053
              |||||||||||||||||||||:|||||||||||||||||||||||||
Db        242 GTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATG 289


RESULT 314
B63737
LOCUS       B63737                  381 bp    DNA     linear   GSS 08-APR-1999
DEFINITION  RPCI11-3D12.TPB RPCI-11 Homo sapiens genomic clone RPCI-11-3D12,
            genomic survey sequence.
ACCESSION   B63737
VERSION     B63737.1  GI:2637727
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-3D12.TVB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7500851"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-3D12"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             3.8%;  Score 283.3;  DB 1;  Length 381;
  Best Local Similarity   82.6%;  Pred. No. 0;
  Matches  314;  Conservative   15;  Mismatches   48;  Indels    3;  Gaps    1;

Qy       4154 CTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCW 4213
              ||||||||: | :||||||||||||:||||||||||||||||||:|||||||:||||:
Db          1 CTAACTTCCTAAGGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGGCT 60

Qy       4214 GTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAA 4273
              | |||||||||||:||||||||:||||||||||| :|||||||||| |||||||||:||
Db         61 GCACAGAAACCTAAAGAGGTGGCAGTCTTACACTGTCGGGGTCATCAGAAAGGAAAGGAA 120

Qy       4274 AGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGAC 4333
              |||||||||:|||:|||: ||||||| :||||||||||:|||||||| |||||||| ||
```

```
Db        121 AGGGAAATAGAAGGGAACAGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAAGGCAAGAT 180

Qy       4334 CCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACC 4393
              ||||||||||||||| ||||  ||   ||| |||||| |||||||||| || ||||||||
Db        181 CCTCCATTAGAAATGTTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACC 240

Qy       4394 AAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTC 4450
              |||||||||||||||| ||| ||| |||||  ||||||||| |||     | ||  |||
Db        241 AAGCCCCAGTACTCAGCTGGAAAAATAGAATACGGAACCTCATGAGAACATACTTCCCTC 300

Qy       4451 CCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGG 4510
              |||||  || || |||||||  |||| || |||||| |||||||||| ||| |||||
Db        301 CCCTCCAGATGGCTAGCCACCAAAGAGGGAAAAATATTTTTGCCTGCAGCTAACCAATAA 360

Qy       4511 AAATTACTTAAAACCCTTCA 4530
              | ||||||||| || | |||
Db        361 ACATTACTTAACACTCGTCA 380


RESULT 315
BG004247
LOCUS       BG004247                 302 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  RC6-GN0070-301100-014-F12 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG004247
VERSION     BG004247.1  GI:12445227
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 302)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            301100-014-F12&t3=2000-11-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 8
            High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .302
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0070"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.7%;  Score 278.6;  DB 1;  Length 302;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  281;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       6577 CTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAA 6636
              |||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          4 CTAGGTCCTGGCATTGGCGGAATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAA 63

Qy       6637 GAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTT 6696
              |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db         64 GAACTAAATGGGGACATGGAACGGGCCGCCGACTCCCTGGCCACCTTGCAAGATCAACTT 123

Qy       6697 AACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGAR 6756
              |||||||||||||||| || |||||:||| ||||||||||||||:||||||||||||:
Db        124 AACTCCCTAGCAGCAGGAGCCCTTCAAAAGCGAAGAGCTTTAGACTTGCTAACCGCTGAA 183

Qy       6757 AGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGA 243

Qy       6817 ATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCT 6875
              |||||||||||||||||:||  || ||:||||||||||||||||: |||||:|||||
Db        244 ATCGTCACTGAGAAAGTTAAAAGAAATCGAGATCGAATACAACGTAGGGCAGAGGAGCT 302


RESULT 316
D29167
LOCUS       D29167                   326 bp    mRNA    linear   EST 24-FEB-2003
DEFINITION  HUMNK203 Human epidermal keratinocyte Homo sapiens cDNA clone 203,
            mRNA sequence.
ACCESSION   D29167
VERSION     D29167.1  GI:599100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Konishi,K., Morishima,Y.-I., Ueda,E., Nonomura,K., Kibe,S.,
            Yamanishi,K. and Yasuno,H.
  TITLE     Cataloging of the genes expressed in human keratinocytes: analysis
            of 607 randomly isolated cDNA sequences
  JOURNAL   Biochem. Biophys. Res. Commun. 202, 976-983 (1994)
   PUBMED   8048971
COMMENT     Contact: Kiyofumi Yamanishi
            Department of Dermatology
            Kyoto Prefectural University of Medicine
            Kamigyo-ku, Kyoto, Kyoto, 602, Japan
            Tel: 75-251-5587
            Fax: 75-251-5587
            PROJECT='human_epidermal_keratinocyte_ESTs'.
FEATURES             Location/Qualifiers
     source          1. .326
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="203"
                     /tissue_type="epidermis"
                     /cell_type="keratinocyte"
                     /clone_lib="Human epidermal keratinocyte"
                     /note="primary culture"

  Query Match             3.7%;  Score 276.6;  DB 1;  Length 326;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  298;  Conservative    7;  Mismatches   16;  Indels   12;  Gaps    1;

Qy       7250 GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAG 7309
              |||||||||||||||||||||||||||||||| ||           |||:|||||||||
Db          6 GGGGACTGAGAGACAGGACTAGCTGGATTTCCCAG-----------TCCCTAAGCCTAG 53

Qy       7310 STGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGAC 7369
              :|||||||||||||||||||||||||||||| ||||||||||||||:|||||||  ||
Db         54 CTGGGAAGGTGACCACATCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAAC 113
```

```
Qy       7370 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAA 7429
             |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||||||
Db        114 CAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAGTAGCCAA 173

Qy       7430 TCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTT 7489
             ||||:|||||||:|| |||||||| |||||||||| |||||||||||||||||| |   ||
Db        174 TCATCTATTGCCTGGGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATT 233

Qy       7490 CGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
             | |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        234 CAAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTGCCTTTGTATGGGAGCTCTGTTTTCA 293

Qy       7550 TGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                |||||||||||||||||||||||||:||||:
Db        294 CTCTATTTCACTCTATTAAATCTTGCAACTGCG 326


RESULT 317
T70359
LOCUS       T70359                   372 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  yd13a03.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:66988 5' similar to SP:ENV2_MOUSE P11370 RETROVIRUS-RELATED
            ENV POLYPROTEIN ;contains PTR7 repetitive element ;, mRNA sequence.
ACCESSION   T70359
VERSION     T70359.1  GI:681507
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 766
            High qality sequence stops: 293 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 766   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 293.
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:463693"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66988"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
```

```
  Query Match             3.7%;  Score 276.4;  DB 1;  Length 372;
  Best Local Similarity   94.5%;  Pred. No. 0;
  Matches  312;  Conservative    4;  Mismatches   10;  Indels    4;  Gaps    3;

Qy       6928 CCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTT 6987
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTT 60

Qy       6988 RACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCC 7047
              :|||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         61 AACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCC 120

Qy       7048 AAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGA 7107
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        121 AAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGA 180

Qy       7108 TCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTA 7167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTA 240

Qy       7168 CTACGCCCCAATTCAGCAGGAAG--CAGTTAGAGCGGTSGTC-GGCCAACCTCCCCAACA 7224
             |||||||||||||||||||||||      |  |||||||:||| ||||||||||||||||||
Db        241 CTACGCCCCAATTCAGCAGGAAGGCAGTTAGGAGCGGTCGTCGGGCCAACCTCCCCAACA 300

Qy       7225 -GCACTTAGGTTTTCCTGTTGAGATGGGGG 7253
              |||||||||||||||||||     ||||
Db        301 GGCACTTAGGTTTTCCTGTTGGAGATGGGG 330


RESULT 318
B59529/c
LOCUS       B59529                   364 bp    DNA     linear   GSS 20-JUN-1998
DEFINITION  CIT-HSP-2017O20.TR CIT-HSP Homo sapiens genomic clone 2017O20,
            genomic survey sequence.
ACCESSION   B59529
VERSION     B59529.1  GI:2614247
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 364)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .364
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7044070"
                     /db_xref="taxon:9606"
                     /clone="2017O20"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
```

HindIII"

Query Match 3.7%; Score 274.7; DB 1; Length 364;
Best Local Similarity 79.5%; Pred. No. 0;
Matches 290; Conservative 33; Mismatches 41; Indels 1; Gaps 1;

```
Qy       3044 CACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGC 3103
             ||||||| :|||||:|| ||||| | : |||||||  |||:||||||||||:|:|:|
Db        364 CACCCAAGTGCTCTTAAACTTCCTTGGTGCCTGTGGTAACAAGGTTTCCAAACCAAAGGA 305

Qy       3104 TCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGC 3163
             |||:||||| | | |||||||||||||||||||: || :||||||||||||||||||:|||
Db        304 TCAGCTCTGATTATAGCAGGTTAAATACTTAGGGTTAGAATTATCCAAAGGCACCAGGGC 245

Qy       3164 CCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCA 3223
             |||||||||| |:::|| ||||||||| ||||||||| |||||:|||| || | |||||
Db        244 CCTCAGTGAGGCACATATTCAGCCTATAATGGCTTATCGTCATCCCAAAGCCATGAAGCA 185

Qy       3224 ACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCR 3283
             | ||||:|: |||||||||:|||:|||:|||||| ||:||||||| ||||||: |||:
Db        184 ATTAAGAGGCTTCCTTGGCATAACAGGCTTCTGCCAAATATGGATT-CCCAGGTACGGCG 126

Qy       3284 AAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAR 3343
             ||||||||||:|||||: ||||:||||| |||||||||||||||||||||| ||||||:
Db        125 AAATAGCCAGGCCATTATGTACACTAATTTAGGAAACTCAGAAAGCCAATACTCATTTAG 66

Qy       3344 TAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCA 3403
             ||  |||||:|:||||||:::|||||||||||    | || |||||||||| ||| |  |
Db         65 TAGAATGGACACCTGAAGCAGAAGTGGCTTTTTCAGGCCATAAAGAAGGCCCTAACCTTA 6

Qy       3404 AGYCC 3408
                :||
Db           5 GCCCC 1
```

RESULT 319
BI056590
LOCUS BI056590 317 bp mRNA linear EST 15-JUN-2001
DEFINITION RC5-GN0282-210201-013-D07 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI056590
VERSION BI056590.1 GI:14464120
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 317)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-GN0282-210201-013-D07&t3=2001-02-21&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 61
High quality sequence stop: 317.

FEATURES Location/Qualifiers
source 1. .317
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0282"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 3.7%; Score 274.4; DB 1; Length 317;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 275; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy       5991 CSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 6050
             |: || |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         36 CCAGGCACATGGCACCTCTAGCCTCTACAAAGGACTAGATCTCTCAAAACTACATGAAAC 95

Qy       6051 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGA 6110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGA 155

Qy       6111 GGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATA 6170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        156 GGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATA 215

Qy       6171 TGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTC 6230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTC 275

Qy       6231 CGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             |||||||||||||||| ||||||||||:|||||||||||||
Db        276 CGTTTTAGTAGGACCTTTTGTTTCCAATCTGGAAATAACCCA 317
```

RESULT 320
BM721661
LOCUS BM721661 353 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO1-aic-j-05-0-UI.r1 UI-E-EO1 Homo sapiens cDNA clone UI-E-EO1-aic-j-05-0-UI 5', mRNA sequence.
ACCESSION BM721661
VERSION BM721661.1 GI:19041724
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 353)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA

```
           sequence: 1-328, >LTR17#LTR/Retroviral
           Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .353
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EO1-aic-j-05-0-UI"
                     /tissue_type="fetal eye"
                     /dev_stage="fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EO1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EO1 is a normalized cDNA library containing the
                     following tissue(s): fetal eye. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     CGCGTATACC. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI)."

  Query Match             3.6%;  Score 270.5;  DB 1;  Length 353;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  293;  Conservative    6;  Mismatches   29;  Indels    1;  Gaps    1;

Qy        244 TGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCAT 303
              ||||||||||||||||| |:| |||||||||| |||| ||||||||||||||||||||||
Db          1 TGTGCTCCTGATCCAGCAAGGTGCCCATTGCCAGTCCCGATTGGGCTAAAGGCTTGCCAT 60

Qy        304 TGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGG 363
              ||| |||||||||||||||||||||||| | ||||| |||||||||||||| ||||||||
Db         61 TGTTCCTGCACGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTGAACACTAGTCACTGGG 120

Qy        364 TTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGG 423
              ||| | |||||||||| ||||||||||||||||||||:| |||||||||| |||||||||
Db        121 TTCTACGGTTCTCTTCCATGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGG 180

Qy        424 CCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGAR 483
              || |||||||||||||||||||||||||:||| |||| |||||||||||||| ||||:
Db        181 CCTAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAG 239

Qy        484 GCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTG 543
              |||||||||||||||||||| ||  | |||:||||||||||| || ||||||||||
Db        240 GCTTGCCACCATCTTGGAAGCGACCCACCACCATCTTGGAAGTGGCTCGCCACCATCTTG 299

Qy        544 GGAGCTCTGTGAGCAAGGACCCCCCGGTR 572
              ||||||||||||||| ||||||| |||:
Db        300 GGAGCTCTGTGAGCANGGACCCCTGGGTA 328


RESULT 321
BX388919/c
LOCUS       BX388919                320 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX388919 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI022YJ18 3-PRIME, mRNA sequence.
ACCESSION   BX388919
VERSION     BX388919.2  GI:46845083
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30461047.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4215.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAB005ZC06_CS00428_3&c=4215.r

FEATURES             Location/Qualifiers
     source          1. .320
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI022YJ18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

  Query Match             3.6%;  Score 270;  DB 1;  Length 320;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches  276;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       6641 TAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACT 6700
              |||   ||| ||||  | | |  | ||||   |||| |||||| ||| |||| |||||||
Db        320 TAATAAGGGGCATGAGATGAGAGGACGACAGGCTGGACACCTTACAAAATCATCTTAACT 261

Qy       6701 CCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAG 6760
              |||||||||||||||||||||:||||||||||||||||||:||||||||||||:||||
Db        260 CCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAG 201

Qy       6761 GGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCG 6820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCG 141

Qy       6821 TCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAA 6880
              |||||||||||||:|||||||||:||||||||||||||||:||||||:||||||||||
Db        140 TCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAA 81

Qy       6881 ACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGAC 6940
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db          80 ACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGAC 21

Qy       6941 CT 6942
              ||
Db          20 CT 19


RESULT 322
R31587/c
LOCUS       R31587                  373 bp    mRNA    linear   EST 28-APR-1995
DEFINITION  yh76e03.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:135676 3', mRNA sequence.
ACCESSION   R31587
VERSION     R31587.1  GI:787430
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
```

```
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 597
            High quality sequence stops: 235
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 597   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 235.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:541639"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:135676"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             3.6%;  Score 267.7;  DB 1;  Length 373;
  Best Local Similarity   89.4%;  Pred. No. 0;
  Matches  329;  Conservative    6;  Mismatches   26;  Indels    7;  Gaps    5;

Qy        772 CYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATT 831
              |:| |  |  | ||  |   |  ||||||||||||||||||||| ||||||||||||||
Db        366 CTTTTGGTAGAAAAGGGCANATGGGAGTGAAGTGCCATAAGTANAAACTTTCTTTTCATT 307

Qy        832 AAGAGACAACTCACAATTA-TGTAAAAAGTGTGATTTATGCCC--TACAGGAAGCCTTCA 888
              |||||||||||||||||| |||||||||||||||||||||      ||||||||||||
Db        306 AAGAGACAACTCACAATTATTGTAAAAAGTGTGATTTATGCCCCTACAGGGAAGCCTTCA 247

Qy        889 GAGTCTACCTCCC--TATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCC 946
              |||||||||||||     |||||||||||||||||||||||::|||||||||||||||
Db        246 GAGTCTACCTCCCCTATCCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCC 187

Qy        947 TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGC 1006
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGC 127

Qy       1007 CAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCA 1066
              |||||||||||||||||||||||| |||||||||||||||||||||||| |||||||||
Db        126 CAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTC-GCCCAGCCA 69

Qy       1067 GAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAAT 1126
              ||||||||||:||||::|||||||||||||||||||||||||||||||||||||     |
Db          68 GAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGNNNNNT 9

Qy       1127 TCTCAGAT 1134
              ||||||||
Db           8 TCTCAGAT 1


RESULT 323
AA868777/c
LOCUS       AA868777                345 bp    mRNA    linear   EST 16-MAR-1998
DEFINITION  ak52h07.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1409629
            3' similar to contains PTR7.t2 PTR7 repetitive element ;, mRNA
            sequence.
ACCESSION   AA868777
VERSION     AA868777.1  GI:2964222
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 324.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1409629"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares_testis_NHT"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from mRNA obtained from
                     Clontech Laboratories, Inc., and primed with  a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "

  Query Match             3.6%;  Score 267.1;  DB 1;  Length 345;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  289;  Conservative    6;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       7230 TAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGA 7289
              | |||||||||||||||| |||||||||||||||||||||||| ||||||||||  |
Db        345 TGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGCTGAATTTCCTAGGCCAA 286

Qy       7290 YTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTG 7349
              :|||||||||:|||| ||||:|||||||||||||| |||||||||||||||||||||| |||
Db        285 CTAAGAATCCCTAAGGCTAGCTGGGAAGGTGACCACGTCCACCTTTAAACACGGGGCTTG 226

Qy       7350 CAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACA 7409
              ||||||||:||||| |||||||||||||||||| |||||||||||||||||||||| |||
Db        225 CAACTTAGCTCACATCTGACCAATCAGAGAGCTCAGTAAAATGCTAATTAGGCAAAAACA 166

Qy       7410 GGAGGTAAAGAAATAGCCAATCAT-YTATTGCMTGAGAGCACAGCAGGAGGGACAATGAT 7468
              ||| |||||||||||||||||| :|||  |:||||||||||   ||| ||||||||||
Db        165 GGACGTAAAGAAATAGCCAATCATCTTATCACCTGAGAGCACAATGGGAAGGACAATGAT 106

Qy       7469 CGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTT 7528
              |||||| |||||||||  |||||||| ||||| || |||| ||||||||||||||||||
```

```
Db        105 CGGGATGTAAACCCAAGCATTCGAGCCAGCAACAGCTACCCTCTTTGGGTCCCCTCCCTT 46

Qy       7529 TGTATGGGAGCTCTGTTTTCA 7549
              ||||||||||||||||| ||||
Db         45 TGTATGGGAGCTCTGTCTTCA 25


RESULT 324
AQ001683/c
LOCUS       AQ001683                370 bp    DNA     linear   GSS 26-JUN-1998
DEFINITION  CIT-HSP-2282F9.TR CIT-HSP Homo sapiens genomic clone 2282F9,
            genomic survey sequence.
ACCESSION   AQ001683
VERSION     AQ001683.1  GI:3026542
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7147548"
                     /db_xref="taxon:9606"
                     /clone="2282F9"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             3.5%;  Score 264.8;  DB 1;  Length 370;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  290;  Conservative    2;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         94 TTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCT 153
              ||||||||| |||||| |||||||||||||| | |||||||||| |  ||||||||||||||
Db        353 TTCACTCTACTAAATCGTGCAACTGCACTCATATGGTCCATGTGTGATACGGCTCGAGCT 294

Qy        154 GAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTC 213
              ||||| |||||||||||||||||||||||||||||||    | | |||  |||||| ||
Db        293 TCGCTTTCGCTCACCGTCCACCACTGCTGTTTGCCACCGTGGTAGACCCACCGCTGGCTT 234

Qy        214 CCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTG 273
              ||||||||| ||||||  |||||| |||||||||||||||||||||:||||||||||
Db        233 CCATCCCTCCGGATCCATCAGGGTATCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTG 174

Qy        274 CCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTG 333
              || ||| | || |||| ||| ||| |||||||| | ||||||||||||||||||| | |
Db        173 CCACTCTCGATCGGGCCAAATGCTCGCCATTGTTCATGCACGGCTAAGTGCCTGGGCTCG 114

Qy        334 TTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCT 393
              | |||| |||||||||||||| |||||||||||| |||||||||| ||||||||||||
Db        113 TCGTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCT 54

Qy        394 TCTAATAKAACTATAACACTTACCACATGGCCCAAGATT 432
              |||||||:| |||||||||| ||  |||||  |||| ||
Db         53 TCTAATAGAGCTATAACACTCACTGCATGGATCAAGCTT 15


RESULT 325
AQ079098/c
LOCUS       AQ079098                328 bp    DNA     linear   GSS 20-AUG-1998
DEFINITION  CIT-HSP-2356G2.TR CIT-HSP Homo sapiens genomic clone 2356G2,
            genomic survey sequence.
ACCESSION   AQ079098
VERSION     AQ079098.1  GI:3440282
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2356G2.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .328
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2356G2"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             3.5%;  Score 260.4;  DB 1;  Length 328;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  284;  Conservative   29;  Mismatches   15;  Indels    4;  Gaps    3;

Qy       3084 AWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARA 3143
              |:|||||||||| |:|:|||||:|||||||||||||| ||||||||||||:||||:|
Db        328 AAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGTTTAAATACTTAGGGCTAAAA 269

Qy       3144 TTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCT 3203
              | |||||||||||||:|||||||||| |||||:::|||| |||||||||||||||||||
Db        268 TCATCCAAAGGCACCAGGGCCCTCAGTCAGGAATGTATCCTGCCTATACTGGCTTATCCT 209

Qy       3204 CATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWA 3263
              ||||:|||||||||||||||||||:|: |||||||||:|||:|||:||||||||||:|
Db        208 CATCCCAAAACCCTAAAGCAACTAAGAGG-TTCCTTGGCATAACAGGTTTCTGCCGAATA 150

Qy       3264 TGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCA 3323
              |||||| || ||||:||||:|||||||||| :|||||:|||||:||||||||||||||||
Db        149 TGGATT-CCTAGGTATGGCAAAATAGCCAGACCATTATATATACACTAATTAAGGAAACTCA 91
```

```
Qy       3324 GAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCC 3383
              |||| |||||||||||||:|||||||||:|:||||| :::|||||||||||||||||||
Db         90 GAAATCCAATACCCATTTAGTAAGATGGACACCTGAATCAGAAGTGGCTTTCCAGGCCCT 31

Qy       3384 TAAAGAAGGCCTTAAACCCAAGYCCCAGTGTT 3415
                ||  | | |||  ||||||||:|||||||||
Db         30 AAAGAAGGCCCT--AACCCAAGCCCCAGTGTT 1


RESULT 326
AA179079
LOCUS       AA179079                352 bp    mRNA    linear   EST 31-DEC-1996
DEFINITION  zp11e01.r1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
            IMAGE:609144 5' similar to SW:POL_AVIRE P03360 POL POLYPROTEIN ;,
            mRNA sequence.
ACCESSION   AA179079
VERSION     AA179079.1  GI:1760431
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 352)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 320.
FEATURES             Location/Qualifiers
     source          1. .352
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:4624403"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:609144"
                     /sex="mixed"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene fetal retina 937202"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Pooled
                     retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

  Query Match             3.5%;  Score 260;  DB 1;  Length 352;
  Best Local Similarity   87.8%;  Pred. No. 0;
  Matches  309;  Conservative   11;  Mismatches   26;  Indels    6;  Gaps    4;

Qy       4414 AGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCAC 4470
              ||||| |||||| ||||||| || |||    ||||| |||||||| ||| || |||||||
Db          1 AGAAATAGAATGAGGAACCTTACAAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCAC 60

Qy       4471 TGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCA 4530
              ||||||||| |||||||||||||||||| ||| ||||||||||||||||||||||||||
Db         61 TGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCA 120

Qy       4531 TCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGG 4590
              |||||||||||||||||||| |||||||||||||:||||||||||||| ||| |||||
Db        121 CCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCATTNTTTNCTGG 180

Qy       4591 ACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAA 4650
              |||||||||||||||||||||||||||:|||:| |||||||||||| |||||||: :| |
Db        181 ACCAGGCCTTTTCAAAACTATCAAGCAGATAGTAAGGGCCTGTGAAATGTGCCAANGAGA 240

Qy       4651 TAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCA-TTACCCTG 4709
              |||||||||||||:||| || |||||||||||||:||||||:|||||||||| |||||| |
Db        241 TAATCCCCTGCCTTATCACCNAGCTCCTTCAGGAGAACAAAGAACAGGCCATTTACCCAG 300

Qy       4710 RARAARACTGGCAACT-GATTTTACCCACAAGCC-CAAACCTCAGGGATTTC 4759
              :|:||:|||||||||| |||||||||||||| ||| |||  ||||||||||||
Db        301 GAGAAGACTGGCAACTAGATTTTACCCACATGCCGCAANTCTCAGGGATTTC 352


RESULT 327
BF943201/c
LOCUS       BF943201                331 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  QV2-NN0045-041000-400-d01 NN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF943201
VERSION     BF943201.1  GI:12360476
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 331)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-NN0045-
            041000-400-d01&t3=2000-10-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 331.
FEATURES             Location/Qualifiers
     source          1. .331
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NN0045"
                     /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

  Query Match             3.5%;  Score 258.8;  DB 1;  Length 331;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  294;  Conservative    6;  Mismatches   28;  Indels    4;  Gaps    2;

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
```

```
                ||||||||||||| |||||||||||  ||   ||| ||||||  |||||||||| ||| |
Db        331 CAGGACCCTCCATAAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCAG 272

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGT 4444
                ||||||||||||||| |||||| |||||| |||||||||||||||| |||    |||
Db        271 AAAACCAAGCCCCAGTATTCAGCAAGAGAAAGAGAATGGGGAACCTCACAAGGACATAGT 212

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
                || |||||||| ||| || ||||||||||||||||||||||||||||| ||||| ||| |
Db        211 TTCCTCCCCTCAGGATGGCTAGCCACTGAAGAAGGGAAAATACTTTTGTCTGCAGCTAAC 152

Qy       4505 CAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCAT 4564
                | |||||||||| || |||||||||| |||||||||||||||||| ||||||||||||
Db        151 CCATGGAAATTA-TTTAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCAT 93

Qy       4565 CARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTC 4624
                ||:|||||||||| ||||||||||||||||||||||||||||||||||||:|||:||
Db          92 CAGATGGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTC 33

Qy       4625 AGGGCCTGTGAAKTGTGCCARARAAATAATCC 4656
                ||||||||||||:|||||||:|:|||||||||
Db          32 AGGGCCTGTGAAGTGTGCCAAAGAAATAATCC 1


RESULT 328
AA774109/c
LOCUS       AA774109                330 bp    mRNA    linear   EST 29-JAN-1998
DEFINITION  ac36d03.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:858533 3', mRNA sequence.
ACCESSION   AA774109
VERSION     AA774109.1  GI:2825998
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 330)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 311.
FEATURES             Location/Qualifiers
     source          1. .330
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:858533"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3'"

  Query Match             3.5%;  Score 258.7;  DB 1;  Length 330;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  296;  Conservative    6;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7251 GGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGS 7310
                |||||||||||||||||||||||||||||||||||||||| ||:|||||||| :|||||||||:
Db        330 GGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCTCTAAGCCTAGC 271

Qy       7311 TGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACC 7370
                ||||||||||||  |||||||||||||||||| |||||||||||||:||||||| ||||
Db        270 TGGGAAGGTGACGGCATCCACCTTTAAACACGTGGCTTGCAACTTAGCTCACACCCGACC 211

Qy       7371 AATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAA 7421
                ||||         |||||||||||||||||||||||||| |||| ||||||||||||||||
Db        210 AATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGACAAAAACAGGAGGTAAAGAA 151

Qy       7422 ATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACC 7481
                ||||||||||||:||| || |||||||||||| |||||||||| ||||||||||||||||
Db        150 ATAGCCAATCATCTATCGCTTGAGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACC 91

Qy       7482 CAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTC 7541
                || |  |||||||||||||| || |||| ||||||||||||||||||||||||||||||||
Db          90 CAGGCATTCGAGCCGGCAACCGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCTC 31

Qy       7542 TGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
                |||           ||||||||||||||||||:||||
Db          30 TGT----------ATTCACTCTATTAAATCTTGCAACTGC 1


RESULT 329
AQ030065
LOCUS       AQ030065                333 bp    DNA     linear   GSS 14-APR-1999
DEFINITION  RPCI11-39G24.TV RPCI-11 Homo sapiens genomic clone RPCI-11-39G24,
            genomic survey sequence.
ACCESSION   AQ030065
VERSION     AQ030065.1  GI:3274029
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-39G24.TP
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7514759"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-39G24"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
```

```
  Query Match             3.4%;  Score 258.3;  DB 1;  Length 333;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  280;  Conservative   15;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4162 YGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAA 4221
              :| |:||||||||||||:|||||||||||||||||||:|||||||:||| :||||||||
Db          1 CGCGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTGTACAGAA 60

Qy       4222 ACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAAT 4281
              |||| :||||||||:|||||||||||||:|||||||||  |||||||||:||||||||||
Db         61 ACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAAGGGAAAT 120

Qy       4282 ASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATT 4341
              |:|||:|||:|||||||| :||||||||||:||||||||  ||||||||||||||||||||
Db        121 AGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACCCTCCATT 180

Qy       4342 AGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCA 4401
              |||||||||||  ||   | | |||||| | |||||||| || |||||||||| || ||
Db        181 AGAAATGCTTATAGAAAGACACCTAGTATGGNGTAATCCCCTCTGGGAAACCAACCCTCA 240

Qy       4402 GTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCCCCTCGGG 4458
              |||||||||| ||||| ||||||||||||||| |||     ||||| || |||||  |
Db        241 GTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTNCCCTCAAG 300

Qy       4459 ACGGTTAGCCACTGAAGAAGGGAAAATACTTTT 4491
              | || ||||||||||||||  |||||||| |
Db        301 ATGGCTAGCCACTGAAGAAGAAAAAATACTTCT 333


RESULT 330
T52094
LOCUS       T52094                  344 bp    mRNA    linear   EST 06-FEB-1995
DEFINITION  yb10h02.r1 Stratagene placenta (#937225) Homo sapiens cDNA clone
            IMAGE:70803 5' similar to contains PTR7 repetitive element, mRNA
            sequence.
ACCESSION   T52094
VERSION     T52094.1  GI:653954
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 344)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 598
            High qality sequence stops: 283 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 598   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 283.
FEATURES             Location/Qualifiers
     source          1. .344
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:491700"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:70803"
                     /sex="male"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene placenta (#937225)"
                     /note="Organ: placenta; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP
                     XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

  Query Match             3.4%;  Score 257.5;  DB 1;  Length 344;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  286;  Conservative    5;  Mismatches    0;  Indels    3;  Gaps    3;

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 TAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACAT 60

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 120

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 AGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 180

Qy       7245 AGATGGGGGACTGA-GAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAA 7303
              |||||||||||||| |||||||||||||||||||||||||||||:|||||||||:|||
Db        181 AGATGGGGGACTGAGGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAA 240

Qy       7304 GCCTAGST-GGGAAGGTGACCACATCCACC-TTTAAACACGGGGCTTGCAACTT 7355
              ||||||:| |||||||||||||||||||| ||||||||||||||||||||||||
Db        241 GCCTAGCTGGGGAAGGTGACCACATCCACCTTTTAAACACGGGGCTTGCAACTT 294


RESULT 331
AG266374/c
LOCUS       AG266374                327 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  Homo sapiens DNA, clone: DN0577, genomic survey sequence.
ACCESSION   AG266374
VERSION     AG266374.1  GI:37573524
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kato,M., Onishi,Y., Wada-Kiyama,Y., Abe,T., Ikemura,T., Kogan,S.,
            Bolshoy,A., Trifonov,E.N. and Kiyama,R.
  TITLE     Dinucleosome DNA of human K562 cells: experimental and
            computational characterizations
  JOURNAL   J. Mol. Biol. 332 (1), 111-125 (2003)
   PUBMED   12946351
REFERENCE   2  (bases 1 to 327)
  AUTHORS   Kato,M., Kiyama,R., Onishi,Y., Ikemura,T. and Abe,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2003) Megumi Kato, National Institute of Advanced
            Industrial Science and Technology, Gene Dynamics Team, Reseach
            Center for Glycoscience; 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566,
            Japan (E-mail:katou-megu@aist.go.jp, Tel:81-298-61-6189,
            Fax:81-298-61-6190)
FEATURES             Location/Qualifiers
     source          1. .327
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="DN0577"
                     /cell_line="K562"

  Query Match             3.4%;  Score 256.5;  DB 1;  Length 327;
  Best Local Similarity   87.2%;  Pred. No. 0;
```

```
  Matches  285;  Conservative    4;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        186 GCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTG 245
              ||| ||| || | ||| ||| ||| || |||||||||  ||||| ||||||||||:||||
Db        327 GCCGCCATCGGAGACCCGCCACTGCCTTCCATCCCTCCAGATCCGGCAGGGTGTCTGCTG 268

Qy        246 TGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTG 305
              |||||| |||||||| |:||||||||||| |||||  ||  ||||||||||| |||||||
Db        267 TGCTCCCGATCCAGCAAGGCGCCCATTGCTGCTCCTGATCTGGCTAAAGGCTCGCCATTG 208

Qy        306 TNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTT 365
              | ||||||| |||||||||| ||||||| |||||||||||||||||||| ||||||||||
Db        207 TTCCTGCACAGCTAAGTGCCCGGGTTTGCCCTAATTGAGCTGAACACTAGTCACTGGGTT 148

Qy        366 CCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCC 425
              |||||||||||||||||||||||||||||||||:| |||||||||| ||| |||||||
Db        147 CCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCC 88

Qy        426 CAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGC 485
              |||||||||||| ||||||||||||||:||| |||| |||||||||||| || |:||
Db         87 CAAGATTCCATTCTTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACAAGGC 29

Qy        486 TTGCCACCATCTTGGAAGCGGCCTGCT 512
              |||| |||||||| ||||| ||| |||
Db         28 TTGCGACCATCTTTGAAGCAGCCCGCT 2


RESULT 332
AI797629/c
LOCUS       AI797629                332 bp    mRNA    linear   EST 18-DEC-1999
DEFINITION  we82a08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2347574 3', mRNA sequence.
ACCESSION   AI797629
VERSION     AI797629.1  GI:5363101
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 808   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 323.
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2347574"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-PacI; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (fetal lung NbHL19W, testis NHT, and
                     B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made
                     in vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "

  Query Match             3.4%;  Score 254.3;  DB 1;  Length 332;

  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  290;  Conservative    8;  Mismatches   25;  Indels   11;  Gaps    2;

Qy       7258 AGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAG 7317
              |||||||||||||||||||||||||||||| ||||:||| |||||:|||||||||:|||||||
Db        332 AGAGACAGGACTAGCTGGATTTCCTAGACTGACTAACAATCCCTAAGCCTAGCTGGGAAG 273

Qy       7318 GTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC--- 7374
              |||||| | |||||||||||||||||| ||||||||||||:|||||||||||||||
Db        272 GTGACCGCTTCCACCTTTAAACACGGGACTTGCAACTTAGCTCACACCTGACCAATCAGG 213

Qy       7375 ------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCA 7428
                    |||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        212 TAGGAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAGGAGGTAAAGAAATAGCCA 153

Qy       7429 ATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYT 7488
              |||||:|||  |:||||| ||||   |||||||||||||| ||||||||||||| || |
Db        152 ATCATCTATCACCTGAGAACACAATGGGAGGGACAATGATCAGGATATAAACCCAGGTAT 93

Qy       7489 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTC 7548
              |||||| |||||||| |||| |||||||||||||||||||||||| |||  ||||||||
Db         92 TCGAGCTGGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGCGAG--CTGTTTTC 35

Qy       7549 ATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |  ||||||||||||||||||||   |:||||:
Db         34 ACTCTATTTCACTCTATTAAATCTAAAAACTGCA 1


RESULT 333
AQ058742/c
LOCUS       AQ058742                351 bp    DNA     linear   GSS 30-JUL-1998
DEFINITION  CIT-HSP-2343I24.TR CIT-HSP Homo sapiens genomic clone 2343I24,
            genomic survey sequence.
ACCESSION   AQ058742
VERSION     AQ058742.1  GI:3355268
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 351)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2343I24.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .351
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2343I24"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
```

```
  Query Match             3.4%;  Score 254.3;  DB 1;  Length 351;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  271;  Conservative   30;  Mismatches   22;  Indels    3;  Gaps    2;

Qy       3097 SARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCA 3156
              :|:|:|||||:||||||||||||||||||||||| ||||||:||||:|||||||||||||||
Db        351 CAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATGCTTAGGGCTAAAATTATCCAAAGGCA 292

Qy       3157 CCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCC 3216
              |||:|||||||||||||||||:::|||||||| ||| |||||||||||||||:|||||||
Db        291 CCAGGGCCCTCAGTGAGGAACGTATCCAGCTTATTCTGGCTTATCCTCATCCCAAAACCA 232

Qy       3217 TAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGG 3276
              |||||||||||||:|::||||||||||:|||:|||:|||||| |||: || ||| ||||||
Db        231 TAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCTGAATGTGAATT-CCCAGG 173

Qy       3277 TWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              |:||| :|||||| ||| :|||||:||||||||||||||||||||||| ||||| |
Db        172 TATGGTGAAATAGACAGACCATTATATACACTAATTAAGGAAACTCAGAAAGTCAATATC 113

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              ||||||:|||||||||:|:||||| :::|||||||||||||||||   ||  | | ||
Db        112 CATTTAGTAAGATGGACACCTGAAACAGAAGTGGCTTTCCAGGCCC--TAAAGAAGTCTC 55

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTG 3422
              |||||||| | ||||||||||:|||
Db          54 TAACCCAAGACTCAGTGTTAAGCTTG 29
```

```
RESULT 334
AQ116788
LOCUS       AQ116788                349 bp    DNA     linear   GSS 22-SEP-1998
DEFINITION  HS_2234_B1_B12_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2234 Col=23 Row=D, genomic survey
            sequence.
ACCESSION   AQ116788
VERSION     AQ116788.1  GI:3494579
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 349)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2234  row: D  column: 23
            Class: BAC ends
            High quality sequence stop: 349.
FEATURES             Location/Qualifiers
     source          1. .349
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2234 Col=23 Row=D"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
```

```
  Query Match             3.4%;  Score 252.4;  DB 1;  Length 349;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  292;  Conservative   29;  Mismatches   21;  Indels   10;  Gaps    4;

Qy       3424 CAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTC 3483
              | ||:|| || || | ||:||:|||::|||||| |||||||||:||:||||||:||||||
Db          1 CCACAGGACATGATTCTTCTTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTC 59

Qy       3484 CTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGAT 3543
              |||||||||:|||:| |||:|||||||||||:: || ::||||||:||| ||||:||||
Db         60 CTTACACAGGTCCGAAGGATGAGCTTGCAACCCGAGGTATACCTGACTAA-GAAATTGAT 118

Qy       3544 GTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTA 3603
              ||||||| |||||||| :|:|||||| |||: |||||||| |||||||||| |::|||||
Db         119 GTAGTGGTAAAGGGTTAGCCTCATTGATTACAGGTAGTGGAGGCAGTAGCACTCTTAGTA 178

Qy       3604 TCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTG 3663
              |||||||||||||||||| ||| ||:|||||||||||||||||||      ||||| |||
Db         179 TCTGAAGCAGTTAAAATAACACATGGAAGAGATCTTACTGTGTGGA-----CATGACGTG 233

Qy       3664 AAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRT 3723
              ||::|||||||||||||||| ||||||   |||||||||||||:||||||||||||:|
Db         234 AACGGCATACTCACTGCTAAAGAAGACTT---GCTGTCAGACAACCGTTTACTTAAATAT 290

Qy       3724 CAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTC 3775
              || |||||||||||||:|||||  |||||:|||||||||||||||||||||
Db         291 CAAGCTCTATTACTTGAAGGGCCATNGCTGCGACTGTGCACTTGTGCAACTC 342
```

```
RESULT 335
H04137/c
LOCUS       H04137                  293 bp    mRNA    linear   EST 20-JUN-1995
DEFINITION  yj45a09.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:151672 3', mRNA sequence.
ACCESSION   H04137
VERSION     H04137.1  GI:867070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 293)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 680
            High qality sequence stops: 199
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 680   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 199.
FEATURES             Location/Qualifiers
     source          1. .293
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:563796"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:151672"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
```

/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta Nb2HP"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I --oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. "

Query Match 3.4%; Score 252; DB 1; Length 293;
Best Local Similarity 92.8%; Pred. No. 0;
Matches 270; Conservative 7; Mismatches 12; Indels 2; Gaps 2;

```
Qy       7294 GAATCCYTAAGCCTAGSTGGGAAGGTGA-CCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             ||| ||:||||||||:|||||||||||| |||||||||||||||| ||||||||||||
Db        293 GAANCCCTAAGCCTAGCTGGGAAGGTGACCCACATCCACCTTTAAANACGGGGCTTGCAA 234

Qy       7353 CTTAGYTCACACCTGA-CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGG 7411
             |||||:|||||||||| ||||||||||||||||||||||||||||||||| ||||||||
Db        233 CTTAGCTCACACCTGACCCAATCAGAGAGCTCACTAAAATGCTAATTAGGCCAAGACAGG 174

Qy       7412 AGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGG 7471
             ||||||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db        173 AGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGG 114

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGT 7531
             |||||||||||| |:|||||||||| ||||| || ||||||||||||||||||| ||||
Db        113 GATATAAACCCAANTCTTCGAGCCGGNAACGGNAAACCCCTTTGGGTCCCCTCCCNTTGT 54

Qy       7532 ATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||||||||||||||||||||||  ||  |||:
Db         53 ATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTNNCANNTGCA 3
```

RESULT 336
AQ241082/c
LOCUS AQ241082 327 bp DNA linear GSS 21-APR-1999
DEFINITION RPCI11-64J14.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-64J14, genomic survey sequence.
ACCESSION AQ241082
VERSION AQ241082.1 GI:3668666
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 327)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: RPCI11-64J14.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .327

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7524421"
/db_xref="taxon:9606"
/clone="RPCI-11-64J14"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

Query Match 3.4%; Score 251.9; DB 1; Length 327;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 288; Conservative 5; Mismatches 21; Indels 13; Gaps 2;

```
Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA- 7035
             ||||||||||:||||||||||| | |||||||||||||| |||| |||:||||||
Db        327 CCTGTATCTTTAACCTCCTTGTTAAGTATGTCTCTTCCAGAATAGAAGTTGTAAAACTAC 268

Qy       7036 ----------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCC 7084
                      |||||||||||| |||||||||||| |||||||||||||| |||||||||
Db        267 AAATGGTTCTTCAAATGGAGCCCCAGATGCAGTCCATGACTAAGATCTACTGCAGACCCC 208

Qy       7085 TGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGG 7144
             |||||| |||||:||||||| | ||||||||||||||||||||||||||||||||||||
Db        207 TGGACCAGCCTGCTAGCCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGG 148

Qy       7145 AAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT-AGAGCGGT 7203
             |||||||| |||||| ||| ||||||||||||||||||||||||||||||| ||||||||
Db        147 AAATCTCAACTGCACGACCCCTACTACGCCCCAATTCAGCAGGAAGCAGTTAAGAGCGGT 88

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              :|| |||||||||||||||| ||||| |||||||||||||  |||||||||||| |
Db         87 CGTTGGCCAACCTCCCCAACATCACTTGAGTTTTCCTGTTGAGAGTGGGGACTGAGAGGC 28

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAY 7290
             || |||||||||||||||||||||||:
Db         27 AGTACTAGCTGGATTTCCTAGGCTGAC 1
```

RESULT 337
AQ007989/c
LOCUS AQ007989 311 bp DNA linear GSS 27-JUN-1998
DEFINITION CIT-HSP-2292E3.TR CIT-HSP Homo sapiens genomic clone 2292E3, genomic survey sequence.
ACCESSION AQ007989
VERSION AQ007989.1 GI:3085215
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 311)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2292E3.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

```
FEATURES             Location/Qualifiers
     source          1. .311
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7151358"
                     /db_xref="taxon:9606"
                     /clone="2292E3"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             3.3%;  Score 250.2;  DB 1;  Length 311;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  264;  Conservative   29;  Mismatches   17;  Indels    2;  Gaps    2;

Qy       3061 ATTTCCTCGCYACCTGTGGCTACAW-GGTTTCCAAACSARARGCTCARCTCTGCTCACAG 3119
              ||||||||||:|||||||||||||: ||||||||| | |:|:|||||: ||||||||||
Db        311 ATTTCCTCGCCACCTGTGGCTACAAGGGTTTCCAAGCTAAAGGCTCAGATCTGCTCACAG 252

Qy       3120 CAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRY 3179
              ||||||||||||||||:||||:| ||||||||||||||:|||||||||||||||||:::
Db        251 CAGGTTAAATACTTAGGGCTAAAACTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGT 192

Qy       3180 ATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT 3239
              |||||||||||| ||||||||||||||: ||||||| ||| ||||||||:|:: |||||
Db        191 ATCCAGCCTATATTGGCTTATCCTCATCCTAAAACCCGAAAACAACTAAGAGGGCTCCTT 132

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
              |||:| |:|||:||||||| ||:||||||| |||||||:  ||:||||||||||:||||
Db        131 GGCATCACAGGCTTCTGCCAAATATGGATT-CCCAGGTACAGCGAAATAGCCAGGCCATT 73

Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGA 3359
              | |||||:|||||||| |||||||||||||||||||||||||:|||||||||:|:||||
Db          72 ACATACACTAATTAAGAAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGA 13

Qy       3360 AGYMRAAGTGGC 3371
              ||:::|||  ||
Db          12 AGCAGAAGCAGC 1


RESULT 338
BI007745
LOCUS       BI007745                299 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  RC5-RT0055-010201-013-B02 RT0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI007745
VERSION     BI007745.1  GI:14411816
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-RT0055-
            010201-013-B02&t3=2001-02-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 95
            High quality sequence stop: 299.
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="RT0055"
                     /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

  Query Match             3.3%;  Score 247.4;  DB 1;  Length 299;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  266;  Conservative    1;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              | |  | ||||     | |||||||||||||| ||||:|||| | ||||||||| ||||
Db          1 CGGCTGGAAAATACTTAGACCCTATTGAACTTGTCAACCTCGGGTCTTTATAATACAGAT 60

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              |||||| ||||| | |||  |||||||||||| | |||||||||||||| ||||||||||
Db         61 CAGGAGAAGCAGCCCGAATGGGACAAACGGGGTAAAAAAAAAGGCCACTGCTTTAGTCAT 120

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              | |||||||||||| |||||||||| |||||||||||||||||||||||||||||| ||||
Db        121 GGCCCTCAGGCAAGCGGACTTTGGAGTCTCTGGAAAAGGGAAAAGCTGGGCAAATCGAAT 180

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
              |||||||| |||||||||||||| || ||||||||||||||||||||||||||||||||
Db        181 GCCTAATATGGCTTGCTTCCAGTGTGGGCTACAAGGACACTTTAAAAAAGATTGTCCAAG 240

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCC 1973
              |||||||||| ||||| |||||||||||||||| ||||||||||||||||||||||||
Db        241 TAGAAGTAAGCTGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGCCC 299


RESULT 339
BF990710
LOCUS       BF990710                305 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  RC6-GN0070-271000-013-D08 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF990710
VERSION     BF990710.1  GI:12397035
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 305)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
```

```
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-GN0070-
            271000-013-D08&t3=2000-10-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 304.
FEATURES              Location/Qualifiers
     source           1. .305
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="GN0070"
                      /note="Organ: placenta_normal; Vector: puc18; Site_1:
                      SmaI; Site_2: SmaI; A mini-library was made by cloning
                      products derived from  ORESTES PCR (U.S. Letters Patent
                      application No. 196,716 -  Ludwig Institute for Cancer
                      Research) profiles into  the pUC 18 vector. Reverse
                      transcription of tissue mRNA and cDNA amplification were
                      performed under low stringency conditions."

  Query Match             3.3%;  Score 245.8;  DB 1;  Length 305;
  Best Local Similarity   95.1%;  Pred. No. 0;
  Matches  250;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       6572 GTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTAT 6631
              | |  |||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GAGGCCTAGGTCCTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTAT 70

Qy       6632 CTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATC 6691
              |||||||||||||| |||||||||| |||||||||||||||||||| ||||||||||||||
Db         71 CTCAAGAACTAAACGGGGACATGGGACGGGTCGCCGACTCCCTGGCCACCTTGCAAGATC 130

Qy       6692 AACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCG 6751
              |||||| |||||||||||||||| |||||:|||||||||||||||||||:||||||||
Db        131 AACTTATCTCCCTAGCAGCAGTAGCCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCG 190

Qy       6752 CTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAAT 6811
              ||||:|||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        191 CTGAAAGAGGGGGAACCTGTTTATTATTAGGGGAAGAATGCTGTTATTATGTTAATCAAT 250

Qy       6812 CCGGAATCGTCACTGAGAAAGTT 6834
              |||||||||||||||||||||||
Db        251 CCGGAATCGTCACTGAGAAAGTT 273


RESULT 340
R27591/c
LOCUS       R27591                   335 bp    mRNA    linear   EST 25-APR-1995
DEFINITION  yh64g07.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:134556 3', mRNA sequence.
ACCESSION   R27591
VERSION     R27591.1  GI:783726
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 335)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1180
            High quality sequence stops: 253
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1180   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 253.
FEATURES              Location/Qualifiers
     source           1. .335
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:540395"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:134556"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares placenta Nb2HP"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             3.3%;  Score 244.7;  DB 1;  Length 335;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  272;  Conservative    7;  Mismatches    3;  Indels    3;  Gaps    3;

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGA-CCACATCCACCTTTAAACACGGGGCTTGCAACTTA 7356
              ||:||| |||| :|||||||||| ||| ||||||||||||||||||||||||||||||||
Db         284 CCCTAANCCTANCTGGGAAGGTGACCCANATCCACCTTTAAACACGGGGCTTGCAACTTA 225

Qy       7357 GYTCACACCTGA-CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 7415
              |:|||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         224 GCTCACACCTGACCCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 165

Qy       7416 AAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATA 7475
              |||||||||||||||||:|||||:||||||||||||||||||||||||||||||||||||
Db         164 AAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATA 105

Qy       7476 TAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 7535
              ||||||||||:|||||| ||||||||||||||||||||||||||||||||||||||||||
Db         104 TAAACCCAAGTCTTCGAG-CGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 46

Qy       7536 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||||||||||||||||||||||||||||||:|||
Db          45 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 341
BQ353502/c
LOCUS       BQ353502                 334 bp    mRNA    linear   EST 20-MAY-2002
DEFINITION  IL3-HT0617-050900-263-F04 HT0617 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ353502
VERSION     BQ353502.1  GI:21017558
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 334)
```

```
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-HT0617-
            050900-263-F04&t3=2000-09-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 167.
FEATURES             Location/Qualifiers
     source          1. .334
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0617"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

  Query Match             3.3%;  Score 243.5;  DB 1;  Length 334;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  272;  Conservative    5;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              |||||||| ||:|| ||| ||||| ||||  || | ||||||||||||||||||| | ||
Db        334 TGATCCAGGGAAGCACCCGTTGCCACTCCTGATCGCGCTAAAGGCTTGCCATTGTTCTTG 275

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              || ||||||||||||||| || || |||||  |||||||||||| |||||||||||||| ||
Db        274 CATGGCTAAGTGCCTGGTTTCGTCCTAATCAAGCTGAACACTAGTCACTGGGTTCCACGG 215

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||| |||||||||| |||||||||:| |||||||||| |||  | ||||||||||
Db        214 TTCTCTTCCGTGACCCACGACTTCTAATAGAGCTATAACACTCACCGTACGGCCCAAGAT 155

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||| ||||||||||:||| |||| ||||||||||||  ||||:||||||||
Db        154 TCCATTCCTTCGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAGCACGAGGCTTGCCA 96

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCT 551
              |||||||||||| |||||||  ||:| ||||||||||  ||||||||||||||||||||
Db         95 CCATCTTGGAAGTGGCCTGCCGCCATTTTGGAAGTGGCCCACCACCATCTTGGGAGCTCT 36

Qy        552 GTGAGCAAGGACCCC 566
                |||||||||||||
Db         35 CGGAGCAAGGACCCC 21


RESULT 342
T87403
LOCUS       T87403                  283 bp    mRNA    linear   EST 17-MAR-1995
DEFINITION  yd89a07.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:115380 5', mRNA sequence.
ACCESSION   T87403
VERSION     T87403.1  GI:715755
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 283)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1256
            High quality sequence stops: 220 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1256   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 220.
FEATURES             Location/Qualifiers
     source          1. .283
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:470997"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:115380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             3.2%;  Score 242.4;  DB 1;  Length 283;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  252;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          9 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGNA 68

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 128

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        129 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 188

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||||:|||||||||||||||||||||||||||   || | | |
Db        189 CGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGGCTAAAGGCTTG 248

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGT 334
              |||||| ||||||  | |||    |||||  | |
Db        249 CATTGTTCCTGCAANGGTAAAGTGCTGGGGGTTT 282
```

```
RESULT 343
CA848155/c
LOCUS       CA848155                292 bp    mRNA    linear   EST 16-DEC-2002
DEFINITION  ip34e12.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6219358 3',
            mRNA sequence.
ACCESSION   CA848155
VERSION     CA848155.1  GI:26998812
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6219358"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

  Query Match             3.2%;  Score 242.2;  DB 1;  Length 292;
  Best Local Similarity   90.6%;  Pred. No. 0;
  Matches  270;  Conservative    7;  Mismatches   15;  Indels    6;  Gaps    2;

Qy       7283 AGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACG 7342
             |||| ||:|||||||||:|||||||||:|||||||||||| ||||||||||||||||||
Db        292 AGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACG 233

Qy       7343 GGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||||||||||:||||||    |||| ||||||| ||||||||||||||||||||
Db        232 GGGCTTGCAACTTAGCTCACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGC 177

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| | |||||||| ||||||||||||||||:|||||:|||||||||||| ||||||||
Db        176 AAAAATAGGAGGTACAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGAC 117

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
```

```
             || |||| |||||||||||||| |  ||||||| |||||||||||||||||||||||||
Db        116 AAGGATCAGGATATAAACCCAGGCATTCGAGCTGGCAACGGCAACCCCCTTTGGGTCCCC 57

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
             |||||||||||||||||||||||||||  || |||||||||||||||||||||:|||
Db         56 TCCCTTTGTATGGGAGCTCTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTG 1


RESULT 344
R37932
LOCUS       R37932                  259 bp    mRNA    linear   EST 04-MAY-1995
DEFINITION  yh97h08.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:137727 5', mRNA sequence.
ACCESSION   R37932
VERSION     R37932.1  GI:795388
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1316
            High quality sequence stops: 170
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1316   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 170.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:544046"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:137727"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo. "

  Query Match             3.2%;  Score 240.1;  DB 1;  Length 259;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches  250;  Conservative    2;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       6408 CTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCC 6467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 CTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCC 65

Qy       6468 YATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAA 6527
```

```
              :|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAA 125

Qy       6528 AAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGG 6587
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        126 AAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGT-CTAGGTGCACTAGGTACTGG 184

Qy       6588 CATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGG 6647
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CATTNGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGG 244

Qy       6648 GGACATGGAACGGGT 6662
              |||||||| ||||||
Db        245 GGACATGGGACGGGT 259


RESULT 345
T87297/c
LOCUS       T87297                 272 bp    mRNA    linear   EST 17-MAR-1995
DEFINITION  yd89a07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:115380 3', mRNA sequence.
ACCESSION   T87297
VERSION     T87297.1  GI:715649
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 272)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1256
            High quality sequence stops: 254 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1256   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 254.
FEATURES             Location/Qualifiers
     source          1. .272
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:470997"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:115380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             3.2%;  Score 239.4;  DB 1;  Length 272;
  Best Local Similarity   95.6%;  Pred. No. 0;

  Matches  258;  Conservative    6;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
              ||| | || |||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        272 GGANGTTGCCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAA 213

Qy       7373 TCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCA 7432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 TCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCA 153

Qy       7433 TYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGA 7492
              |:||||||:|||||||||||||||||||||||||||||||||||||||||||:|||||
Db        152 TCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGA 93

Qy       7493 GCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGC 7552
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         92 -NCGGCAACGGCAA-CCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGC 35

Qy       7553 TATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||||||||||:||||:
Db         34 TATTTCACTCTATTAAATCTTGCAACTGCA 5


RESULT 346
BG014730/c
LOCUS       BG014730               323 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  QV2-GN0208-211200-582-b09 GN0208 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG014730
VERSION     BG014730.1  GI:12466225
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-GN0208-
            211200-582-b09&t3=2000-12-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 323.
FEATURES             Location/Qualifiers
     source          1. .323
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0208"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
```

performed under low stringency conditions."

Query Match 3.2%; Score 239; DB 1; Length 323;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 266; Conservative 5; Mismatches 52; Indels 0; Gaps 0;

```
Qy        203 GCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGA 262
              || ||||||| ||||||||| ||||   |||||||||| |||||||||||||||| ||
Db        323 GCTGCTGACTTCCATCCCTCCGGATGGAGCAGGGTGTCCACTGTGCTCCTGATCCAGTGA 264

Qy        263 RGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGT 322
              :|| ||||||||| ||||| |||||||||||||||||||||||| |||||| ||||||||
Db        263 GGCACCCATTGCCACTCCCGATTGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGT 204

Qy        323 GCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGT 382
              |||||||||  | |||         |||||||| |||||||||| | |||||||||| ||
Db        203 GCCTGGGTTCATCCTATCCAGCTGGAACACTAGTCACTGGGTTCTACGGTTCTCTTCCGT 144

Qy        383 GACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTG 442
              ||||||| ||||||||||:|  ||||||||| ||  ||||||| ||||||||||||||||
Db        143 GACCCACAGCTTCTAATAGAGATATAACACTCACAGCATGGCCGAAGATTCCATTCCTTG 84

Qy        443 GAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAA 502
              |||||| |||||:|||| | | |  |  ||||||| ||||:|||||||| |||| |||||
Db         83 GAATCCATGAGGCCAACAACCCCAGGTCCAGAGAACACGAGGCTTGCCAACATCGTGGAA 24

Qy        503 GCGGCCTGCTACCRTCTTGGAAG 525
              | ||||  | |||:|||||  ||
Db         23 GTGGCCCACCACCATCTTGAGAG 1
```

RESULT 347
R00744/c
LOCUS R00744 323 bp mRNA linear EST 31-MAR-1995
DEFINITION ye74h09.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:123521 3' similar to SP:ENV_AVISN P31796 ENV POLYPROTEIN ;, mRNA sequence.
ACCESSION R00744
VERSION R00744.1 GI:750480
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 323)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 221
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -21m13
High quality sequence stop: 221.
FEATURES Location/Qualifiers
source 1. .323
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:476066"
/db_xref="taxon:9606"
/clone="IMAGE:123521"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

Query Match 3.2%; Score 237.3; DB 1; Length 323;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 298; Conservative 3; Mismatches 15; Indels 7; Gaps 5;

```
Qy       6488 AAGATTTATACAGTTATGTCATATCTAAGCCCCGC---AACAAAAGAGTACCCATTCTTC 6544
              ||||||||||   |||| ||||||  |  ||| |   ||||||||||||||  || ||
Db        323 AAGATTTATACCAGTATGCCATATCCTAAGCCCCCGNAAACAAAAGAGTACCCCATCCTC 264

Qy       6545 CTTTTGTT-ATAGGAGCAGGAGTG-CTAGGTGCACTAGGTACTGGCATTGGC-GGTATCA 6601
              | |||||| |||||||||||||||| |||||||||||||||||||||||||| |||||||
Db        263 CCTTTGTTAATAGGAGCAGGAGTGCCTAGGTGCACTAGGTACTGGCATTGGCNGGTATCA 204

Qy       6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 6661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 144

Qy       6662 TCGCCGACT-CCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 TCGCCGACTCCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 84

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              |:|||||||||||||||||| : |||||||||:|||||||||||||||||||||||||||
Db          83 CAAAATCGAAGAGCTTTAGACCTTCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTA 24

Qy       6781 GGGGAAGAATGCTGTTATTATGT 6803
              ||||||||||||||||||||||||
Db          23 GGGGAAGAATGCTGTTATTATGT 1
```

RESULT 348
BG015482/c
LOCUS BG015482 322 bp mRNA linear EST 24-JAN-2001
DEFINITION QV2-GN0208-261200-584-g07 GN0208 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG015482
VERSION BG015482.1 GI:12467711
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 322)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

```
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-GN0208-
            261200-584-g07&t3=2000-12-26&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 322.
FEATURES              Location/Qualifiers
     source           1. .322
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="GN0208"
                      /note="Organ: placenta_normal; Vector: puc18; Site_1:
                      SmaI; Site_2: SmaI; A mini-library was made by cloning
                      products derived from  ORESTES PCR (U.S. Letters Patent
                      application No. 196,716 -  Ludwig Institute for Cancer
                      Research) profiles into  the pUC 18 vector. Reverse
                      transcription of tissue mRNA and cDNA amplification were
                      performed under low stringency conditions."

  Query Match             3.1%;  Score 232.3;  DB 1;  Length 322;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  268;  Conservative    5;  Mismatches   49;  Indels    1;  Gaps    1;

Qy        203 GCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGA 262
              || ||||||| |||| |||| ||||   |||||||||| |||||||||||||||| ||
Db        322 GCTGCTGACTTCCATTCCTCCGGATGGAGCAGGGTGTCCACTGTGCTCCTGATCCAGTGA 263

Qy        263 RGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGT 322
              :|| |||| |||| ||||| |||||||||||||||||||||||| |||||| ||||||||
Db        262 GGCACCCACTGCCACTCCCGATTGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGT 203

Qy        323 GCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGT 382
              ||| ||||| | |||        |||||||| |||||||||| | ||||||||| ||
Db        202 GCCCGGGTTCATCCTATCCAGCTGGAACACTAGTCACTGGGTTCTACGGTTCTCTTCCGT 143

Qy        383 GACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTG 442
              ||||||| ||||||||||:|  |||||||||| ||  ||||||| ||||||||||||||||
Db        142 GACCCACAGCTTCTAATAGAGATATAACACTCACAGCATGGCCGAAGATTCCATTCCTTG 83

Qy        443 GAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAA 502
              |||||| |||||:|||| ||| |||||||||||| ||||:|||||||| |||| |||||
Db         82 GAATCCATGAGGCCAAC-AACCCCAGGTCAGAGAACACGAGGCTTGCCAACATCGTGGAA 24

Qy        503 GCGGCCTGCTACCRTCTTGGAAG 525
              | ||||  | |||:|||||  ||
Db         23 GTGGCCCACCACCATCTTGAGAG 1


RESULT 349
AQ000990/c
LOCUS       AQ000990                310 bp    DNA     linear   GSS 26-JUN-1998
DEFINITION  CIT-HSP-2283A5.TF CIT-HSP Homo sapiens genomic clone 2283A5,
            genomic survey sequence.
ACCESSION   AQ000990
VERSION     AQ000990.1  GI:3028429
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .310
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7147808"
                      /db_xref="taxon:9606"
                      /clone="2283A5"
                      /sex="Male"
                      /cell_type="Sperm"
                      /clone_lib="CIT-HSP"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII"

  Query Match             3.1%;  Score 230.7;  DB 1;  Length 310;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  266;  Conservative   28;  Mismatches   16;  Indels    5;  Gaps    4;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||||||||||||||||||||||||||:||||:||||||||||||||||:
Db        310 GGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 251

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              ||||||||||||||:::|||||||||||||||||||||||||||:||||||||||||||
Db        250 AGCCCTCAGTGAGGAATATATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAA 191

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::|| |||||:|||:|||:||||||| ||: ||||||| |||||||: |
Db        190 GCAACTAAGAGGGTT-CTTGGCATAACAGGCTTCTGCCAAATGTGGATT-CCCAGGTACG 133

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              | :| |||||||||:| |||:|||||:|||||||||||||||| ||||||||||||||||
Db        132 GTGAGATAGCCAGGCCGTTATATACACTAATTAAGGAAACTCAG-AAGCCAATACCCATT 74

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:||||||||||:|:||||||:::||| |||| |||||||||  ||  | | ||  ||||
Db          73 TAGTAAGATGGACACCTGAAGCAGAAGAGGCTGTCCAGGCCCTAAAGAAGGCCC--AAAC 16

Qy       3401 CCAAGYCCCAGTGTT 3415
              |||||: ||||||||
Db          15 CCAAGCTCCAGTGTT 1


RESULT 350
BF989696/c
LOCUS       BF989696                291 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  MR2-GN0127-231000-004-e11 GN0127 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF989696
VERSION     BF989696.1  GI:12396021
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 291)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
```

```
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR2&t2=MR2-GN0127-
            231000-004-e11&t3=2000-10-23&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 290.
FEATURES             Location/Qualifiers
     source          1. .291
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0127"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             3.1%;  Score 230.4;  DB 1;  Length 291;
  Best Local Similarity   91.7%;  Pred. No. 0;
  Matches  255;  Conservative    7;  Mismatches   14;  Indels    2;  Gaps    2;

Qy       2681 GATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTA 2740
              || || ||||||||||||  ||||| |||||  | |  |||| ||||||||||||||||||||
Db        280 GAACTTAGGATTATCAAT-GAGGCTGTTGTTCCTCTATAGCCAGCTGTACCTAGCCCTTA 222

Qy       2741 TACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGA 2800
              ||||  ||:|||||||||||||||||||||||||||||||||||:||||||||||:||||
Db        221 TACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGA 162

Qy       2801 TGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATAC 2860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATAC 102

Qy       2861 TTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCC 2920
              ||||||||||:||||||||||||||||||:||||||||||||||||||||||||||:||
Db        101 TTCAAACCCAACATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCC 42

Qy       2921 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYC 2958
              |||||||||| |||||||||||||| |||    ||| |
Db         41 CCATCTATTT-GCCAGGCATTAGCCTAAGTTCAGAGNC 5


RESULT 351
DU131223
LOCUS       DU131223                 269 bp    DNA     linear   GSS 02-SEP-2005
DEFINITION  Hsap_9_UR_A04 Hsap Homo sapiens genomic, genomic survey sequence.
ACCESSION   DU131223
VERSION     DU131223.1  GI:74099467
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   Wicker,T., Robertson,J.S., Schulze,S.R., Feltus,F.A., Ivarie,R. and
            Paterson,A.H.
  TITLE     Ultra-rapidly associating DNA forms diverse secondary structures
            with many biological functions
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Paterson AH
            Plant Genome Mapping Laboratory
            University of Georgia, Center for Applied Genetic Technologies
            Riverbend Research Laboratory, Room 162, 110 Riverbend Road,
            Athens, GA 30602 USA
            Tel: 7065830169
            Fax: 7065830160
            Email: paterson@uga.edu
            Plate: Hsap-9  row: A  column: 04
            Class: unknown.
FEATURES             Location/Qualifiers
     source          1. .269
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Hsap"
                     /note="Vector: pBluescript; Site_1: EcoR1; Site_2: Xho1"

  Query Match             3.1%;  Score 230.1;  DB 1;  Length 269;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  240;  Conservative   15;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       3815 TARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAG 3874
              ||:||:||||||| ||||||:|||||||||||||||   | ||||||| |||||:|||
Db          1 TAGAACATAACTGCCAACAAGTAATTTCTCAAACCTAGATCGCTCGAGGGAACCTTTTAG 60

Qy       3875 ARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAA 3934
              |:|||||:|| ||||||||:|||| ||||||||||||||||||||||||||||||||||
Db         61 AGGTTCCCTTAACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAA 119

Qy       3935 AAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAAT 3994
              ||||||||| ||||||:|||||||||||| ||||||||||||||||:|||||| ||||
Db        120 AAAGGACTTTGAAAAGTGGGGTATGCAGTGCTCAGTGATAATGGAATACTTGAAAATAAT 179

Qy       3995 CCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTA 4054
              |||||||||||||||| || |||:||||:|||||||||||||:|||:| ||||||||
Db        180 CCCCTCACTCCAGGAACCAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTA 239

Qy       4055 GAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
              |||||||||||||:||||||||:||||| |
Db        240 GAATTAGGAGAAGGAAAAAGGGTAAATACA 269


RESULT 352
AQ539232/c
LOCUS       AQ539232                 314 bp    DNA     linear   GSS 19-MAY-1999
DEFINITION  RPCI-11-359C2.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-359C2,
            genomic survey sequence.
ACCESSION   AQ539232
VERSION     AQ539232.1  GI:4869806
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 314)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-359C2.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
```

```
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .314
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7637521"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-359C2"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             3.1%;  Score 230;  DB 1;  Length 314;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  265;  Conservative    3;  Mismatches   18;  Indels   12;  Gaps    2;

Qy       6484 GAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTT 6543
              |||| ||||||||||| | |||   || |||||| |||||||||||||||||||||||||
Db        286 GAACTAGATTTATACAATCATG---TACCTAAGCACCGCAACAAAAGAGTACCCATTCTT 230

Qy       6544 CCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACA 6603
              |||||||||||  |||||||||||||||||  ||  |          ||| |||||||||
Db        229 CCTTTTGTTATCAGAGCAGGAGTGCTAGGTGTGCTGTG---------TGGTGGTATCACA 179

Qy       6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 6663
              |||||||||||||||||||||||| ||||||||||||||| ||||||||| |||||
Db        178 ACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTC 119

Qy       6664 GCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRA 6723
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||:|
Db        118 GCCGACTCCCTGGTCACCTTGCAAGATTAACTTAACTCCCTAGCAGCAGTAGTCCTTCAA 59

Qy       6724 AATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAG 6781
              |||||||||||||||||:||||||| | ||:||||||||||||||||||||||||
Db         58 AATCGAAGAGCTTTAGACTTGCTAACCACGGAAAGAGGGGGAACCTGTTTATTTTTAG 1


RESULT 353
BU790206/c
LOCUS       BU790206                 278 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in52a11.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6125493 3',
            mRNA sequence.
ACCESSION   BU790206
VERSION     BU790206.1  GI:23840413
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6125493"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

  Query Match             3.1%;  Score 229.6;  DB 1;  Length 278;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  247;  Conservative    7;  Mismatches   16;  Indels    4;  Gaps    1;

Qy       7179 TTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 7238
              | | ||||||||| |||||||| ||: | |||||||||||||||||||||| |||||||
Db        270 TCCGGCAGGAAGCTGTTAGAGCAGTCATTGGCCAACCTCCCCAACAGCACTTGGGTTTTC 211

Qy       7239 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATC 7298
              |||||||||  |||||||||||||||||||||||||||||||||||| ||:||||||||
Db        210 CTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATC 151

Qy       7299 CYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGY 7358
              |:|||||||||:|||||||||||||| |||||||||||||||||||||||||||||||:
Db        150 CCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAACTTAGC 91

Qy       7359 TCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAA 7418
              ||||||    |||| ||||||| |||||||||||||||||||| | |||||||| |
Db         90 TCACAC----CCAACCAGAGAGTTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTACA 35

Qy       7419 GAAATAGCCAATCATYTATTGCMTGAGAGCACAG 7452
              |||||||||||||||:||||||:||||||||||
Db         34 GAAATAGCCAATCATCTATTGCCTGAGAGCACAG 1


RESULT 354
BQ365242
LOCUS       BQ365242                 244 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC4-GN0039-250800-012-h02 GN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365242
VERSION     BQ365242.1  GI:21040754
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 244)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
```

```
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC4&t2=RC4-GN0039-
            250800-012-h02&t3=2000-08-25&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 27
            High quality sequence stop: 244.
FEATURES              Location/Qualifiers
     source           1. .244
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="GN0039"
                      /note="Organ: placenta_normal; Vector: puc18; Site_1:
                      SmaI; Site_2: SmaI; A mini-library was made by cloning
                      products derived from  ORESTES PCR (U.S. Letters Patent
                      application No. 196,716 -  Ludwig Institute for Cancer
                      Research) profiles into  the pUC 18 vector. Reverse
                      transcription of tissue mRNA and cDNA amplification were
                      performed under low stringency conditions."

  Query Match             3.0%;  Score 223.8;  DB 1;  Length 244;
  Best Local Similarity   94.9%;  Pred. No. 0;
  Matches  225;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
              || |:    |||||||||||| ||||||:|||||||||||||||||||||||||||||
Db          4 ATCCCGCGGCGAATACAACGTAAAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 63

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 123

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        124 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 183

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
              ||||||||:|||||||||||||||||||||||||||||||||||||||||| |||||
Db        184 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCGACCGC 240


RESULT 355
B46293/c
LOCUS       B46293                  268 bp    DNA     linear   GSS 21-OCT-1997
DEFINITION  HS-1063-A2-G08-MR.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 796 Col=16 Row=M, genomic survey
            sequence.
ACCESSION   B46293
VERSION     B46293.1  GI:2551127
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 268)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 796  row: M  column: 16
            Class: BAC ends
            High quality sequence stop: 268.
FEATURES              Location/Qualifiers
     source           1. .268
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=CT 796 Col=16 Row=M"
                      /sex="M"
                      /clone_lib="CIT Human Genomic Sperm Library C"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"

  Query Match             3.0%;  Score 221.9;  DB 1;  Length 268;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  227;  Conservative   25;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3117 CAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAA 3176
              || |||||||||| ||||||:||||:|||||||||||||| ||:|||||||||| |||||
Db        268 CAACAGGTTAAATGCTTAGGGCTAAAATTATCCAAAGGCAGCAGGGCCCTCAGTAAGGAA 209

Qy       3177 YRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTC 3236
              :::|||||||| |||||||||||||||||:|| | |||||||||||||||:|::  |
Db        208 TGTATCCAGCCCATACTGGCTTATCCTCATCCCAGAGCCCTAAAGCAACTAAGAGGGCAC 149

Qy       3237 CTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYC 3296
              ||||||:|||:|||:||||||| ||:||||||| |||||||:| ||:||||||||||:|
Db        148 CTTGGCATAACAGGCTTCTGCCAAATATGGATT-CCCAGGTATCGCGAAATAGCCAGGCC 90

Qy       3297 ATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMC 3356
              || |:|||| :|||||||||||||||||||||||||||||||||:|||||||||:|:|
Db         89 ATCATATACGCTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACC 30

Qy       3357 TGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
              | |||:::|||||||||| ||||| |||||
Db         29 TAAAGCAGAAGTGGCTTCCCAGGACCCTA 1


RESULT 356
BE082150
LOCUS       BE082150                293 bp    mRNA    linear   EST 12-JUN-2000
DEFINITION  QV2-BT0636-060500-197-c09 BT0636 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE082150
VERSION     BE082150.1  GI:8472446
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 293)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
```

```
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV2-BT0636-060
            500-197-c09&t3=2000-05-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 293.
FEATURES              Location/Qualifiers
     source           1. .293
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="BT0636"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."

  Query Match             3.0%;  Score 221.6;  DB 1;  Length 293;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  262;  Conservative    7;  Mismatches   23;  Indels    4;  Gaps    3;

Qy       1084 YYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGAT 1143
              :: ||| ||||||| |||||||| |||| |||| ||||| |||||||||||||:| ||||
Db          2 CCCCTCTCAGACTTGAAGCAAATTAAAATAGACCTAGGTGAATTCTCAGATAACCTTGAT 61

Qy       1144 GGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA 1203
              ||||||||||:||||||||||||||||||||| ||||||||||||||||||||||||||
Db         62 GGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAGATATA 121

Qy       1204 TATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGC 1263
                |||| |||||||||||||| |||||||||||||||||||||||||||| ||||| |||||||
Db        122 -ATGTTACTGCTAAATCAGGCACTAACCCCAAATGAGAGAAGTGCCGCCATAGCTGCAGC 180

Qy       1264 CTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAA 1323
              | |||:|||||| |||||||||  |||||||||| ||||||     |||||||||||| |
Db        181 CCGAGAGTTTGGTGATCTCTGGCGTCTCAGTCAGTTCAATG--ATAGGATGACAACAG-A 237

Qy       1324 GGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATT 1379
              ||||||| |||||||||||||||||||:|||||||||||| ||:|||||||||
Db        238 GGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATT 293


RESULT 357
R31328
LOCUS       R31328                   242 bp    mRNA    linear   EST 28-APR-1995
DEFINITION  yh76e03.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:135676 5' similar to contains MER22 repetitive element ;,
            mRNA sequence.
ACCESSION   R31328
VERSION     R31328.1  GI:787171
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 242)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 597
            High quality sequence stops: 223
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 597   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 223.
FEATURES              Location/Qualifiers
     source           1. .242
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:541639"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:135676"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares placenta Nb2HP"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             2.9%;  Score 220.9;  DB 1;  Length 242;
  Best Local Similarity   95.8%;  Pred. No. 0;
  Matches  226;  Conservative    9;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        490 CACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCT 549
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db          7 CACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCT 66

Qy        550 CTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAAGTGATG 609
              |||||||||||||||||||||:||||||||:||||:::||||||||||:||||||||
Db         67 CTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAAGTGATG 126

Qy        610 GGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAA 669
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||:|||
Db        127 GGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAA 186

Qy        670 TTTGACCCTCAGACACTAA-GAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCT 724
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        187 TTTGACCCTCAGACACTAAGGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCCT 242


RESULT 358
BQ022458/c
LOCUS       BQ022458                 290 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-1-BB1p-akl-d-12-0-UI.s1 NCI_CGAP_Pl6 Homo sapiens cDNA clone
            UI-1-BB1p-akl-d-12-0-UI 3', mRNA sequence.
ACCESSION   BQ022458
VERSION     BQ022458.1  GI:19757737
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 290)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 1-22, >AT_rich#Low_complexity 52-290,
            >HERV17#LTR/Retroviral (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .290
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="UI-1-BB1p-akl-d-12-0-UI"
                      /tissue_type="Placenta"
                      /dev_stage="Full Term"
                      /lab_host="DH10B (Life Technologies)"
                      /clone_lib="NCI_CGAP_Pl6"
                      /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                      a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                      NCI_CGAP_Pl6 is a subtracted cDNA library constructed
                      according to Bonaldo, Lennon and Soares, Genome Research,
                      6:791-806, 1996. First strand cDNA synthesis was primed
                      with an oligo-dT primer containing a Not I site. Double
                      stranded cDNA was ligated to an EcoR I adaptor, digested
                      with Not I, and cloned directionally into pT7T3-Pac
                      vector. The oligonucleotide used to prime the synthesis of
                      first-strand cDNA contains a library tag sequence that is
                      located between the Not I site and the (dT)18 tail. The
                      sequence tags for this library are GA, AGGAA. For
                      additional information, contact: Bento Soares,
                      bento-soares@uiowa.edu
                      TAG_TISSUE=placenta human full term
                      TAG_LIB=UI-1-BB1p
                      TAG_SEQ=AGGAA"

  Query Match              2.9%;  Score 220.8;  DB 1;  Length 290;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches  228;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5360 CTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGT 5419
             |||||||||||| ||||||||||||||:||||||||||||||||||||||||||||||||
Db        290 CTGGGAAGGACCGTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGT 231

Qy       5420 GGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAA 5479
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        230 GGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTACCAAAGGAACCTGAAAA 171

Qy       5480 TCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAA 5539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 TCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAA 111

Qy       5540 ACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATA 5599
             |||||||||||||||||||||||||||||||||||||||||    : | |  ||   |||
Db        110 ACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCTCCCCTCCCTTTTTGTATA 51

Qy       5600 TTTT 5603
              |||
Db         50 ATTT 47


RESULT 359
H13097
LOCUS       H13097                   267 bp    mRNA    linear   EST 27-JUN-1995
DEFINITION  yj10h08.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:148383 5', mRNA sequence.
ACCESSION   H13097
VERSION     H13097.1  GI:877917
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1370
            High quality sequence stops: 202
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1370   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 202.
FEATURES              Location/Qualifiers
     source           1. .267
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:560141"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:148383"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares placenta Nb2HP"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match              2.9%;  Score 219.8;  DB 1;  Length 267;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches  239;  Conservative    2;  Mismatches   16;  Indels    4;  Gaps    1;

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 AGCAATACTACATACACAACCAACTCCCAATGCNTCAGGTGGGTAACTCCTCCCACACAA 60

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 120

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             |||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db         121 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 180

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCC----GCAACAAAAGAGTACC 6536
             |||||||||||||||||||||||||||||||||||||||| |||     ||| |  ||    ||
Db         181 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGGCCCCGCAACAAAAGGAGTACCCC 240

Qy       6537 CATTCTTCCTTTTGTTATAGG 6557
              |  | ||||||||||| |||
Db         241 ATTCTTNCCTTTTGTTACAGG 261


RESULT 360
```

```
AQ264399/c
LOCUS       AQ264399                239 bp    DNA     linear   GSS 27-OCT-1998
DEFINITION  CITBI-E1-2507P3.TR CITBI-E1 Homo sapiens genomic clone 2507P3,
            genomic survey sequence.
ACCESSION   AQ264399
VERSION     AQ264399.1  GI:3791153
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 239)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2507P3.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .239
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2507P3"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

  Query Match             2.9%;  Score 214.4;  DB 1;  Length 239;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  212;  Conservative   15;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       3908 TATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTC 3967
             |||||||||||| |||||||||||||||||||||||||||||||: ||||||||||||||||
Db        239 TATACTGATGGCAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGCAGGGTATGCAGTGGTC 180

Qy       3968 AGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCA 4027
             |||||||||||||||:|||||||| |||||||||||||||||||||||:  ||: ||
Db        179 AGTGATAATGGAATACTTGAAAGTTATCCCCTCACTCCAGGAACTAGTGCTCCACTGACA 120

Qy       4028 GAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATAT 4087
             |||||||||||:|||:|  |||||||||||||||||||:||||||||:||||||||||
Db        119 GAACTAATAGCCCTCACTCAGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATAT 60

Qy       4088 ACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAG 4146
             |||||||||:|: |||||:|||| ||| |||||||||||::|||| |||:|:||||||
Db         59 ACAGACTCTAAGAATGCTTACCTTGTCCTCCATGCCCATGCAGCAACATGGAGAGAAAG 1


RESULT 361
BQ365241
LOCUS       BQ365241                238 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC4-GN0039-250800-012-e11 GN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ365241
VERSION     BQ365241.1  GI:21040753
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 238)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC4&t2=RC4-GN0039-
            250800-012-e11&t3=2000-08-25&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 238.
FEATURES             Location/Qualifiers
     source          1. .238
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0039"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             2.9%;  Score 214.1;  DB 1;  Length 238;
  Best Local Similarity   96.6%;  Pred. No. 0;
  Matches  224;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       6847 GATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG-CCTCCTCAG 6905
             | |||||||||||||| |||||:|||||||||||||||||||||||| |||||||||
Db          3 GGTCGAATACAACGTAAAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGACCTCCTCAG 62

Qy       6906 CCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACT 6965
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACT 122

Qy       6966 CCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGC 7025
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        123 CCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGC 182

Qy       7026 TGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGC 7077
             |||:|||||||||||||||||||||||||||||||||||||||||||| |||||
Db        183 TGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCGACCGC 234


RESULT 362
H13302/c
LOCUS       H13302                  239 bp    mRNA    linear   EST 27-JUN-1995
DEFINITION  yj10h08.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:148383 3', mRNA sequence.
ACCESSION   H13302
VERSION     H13302.1  GI:878122
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 239)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1370
            High quality sequence stops: 129
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1370   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 129.
FEATURES             Location/Qualifiers
     source          1. .239
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:560141"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:148383"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "

  Query Match             2.9%;  Score 213.9;  DB 1;  Length 239;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  225;  Conservative    5;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATG-CTAATTAGGC 7402
             |||||||| |||||:||| ||||| |||||||||||||||||||||||| ||||||||||
Db        239 GGCTTGCACCTTAGCTCAAACCTGCCCAATCAGAGAGCTCACTAAAATGCCTAATTAGGC 180

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             |||||||||||||||||||||||||||||:||||||:|||||||||||||||||||||||
Db        179 AAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGAC 120

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||||:|||||  ||||||||| || ||||||||||||||||||
Db        119 AATGATCGGGATATAAACCCAAGTCTTCGAAGCGGCAACGGNAAACCCCTTTGGGTCCCC 60

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||| ||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          59 TCCCNTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 1


RESULT 363
R01948
LOCUS       R01948                   294 bp    mRNA    linear   EST 31-MAR-1995
DEFINITION  ye85b02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:124491 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
            POLYPROTEIN - ;, mRNA sequence.
ACCESSION   R01948
VERSION     R01948.1  GI:751684
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1118
            High quality sequence stops: 209 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1118   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 209.
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:477036"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:124491"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

  Query Match             2.8%;  Score 209.3;  DB 1;  Length 294;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  244;  Conservative    6;  Mismatches   31;  Indels    3;  Gaps    2;

Qy       6641 TAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACT 6700
             ||||||| |||||||||| |||| | ||||||||||||||||||||||||||||||||||
Db          2 TAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACT 61

Qy       6701 CCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAG 6760
             ||||||||||||||||||||:||||||||||||||||||:|||||||||   |:||||
Db         62 CCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAG 121

Qy       6761 GGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCG 6820
             |||||||||||||||||||| ||||||||||||||||||||||||||||||| |||| |
Db        122 GGGGAACCTGTTTATTTTTAGGAGAAGAATGCTGTTATTATGTTAATCAATCCAGAATTG 181

Qy       6821 TCACTGAGAAAGTTRAA--GAAATTCSAGATCGAATACAACGTAKAGCAGA-RGAGCTTC 6877
             |||||||| ||  | ||  |||||||:||||||||||||| |||:|||||| :|||||||
Db        182 TCACTGAGGAAAGTTAAAGGAAATTCGAGATCGAATACAATGTAGAGCAGAGGGAGCTTC 241

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGG 6921
```

```
                 || ||| |  |   ||||||||||    | |||| |  || ||
Db        242 AAANCACCGACCATGGGGGCCTCCTTNAGCCATGGGTTNCCCGG 285


RESULT 364
B33508/c
LOCUS       B33508                   289 bp    DNA     linear   GSS 17-OCT-1997
DEFINITION  HS-1022-B2-D12-MR.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 801 Col=24 Row=H, genomic survey
            sequence.
ACCESSION   B33508
VERSION     B33508.1  GI:2532877
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 801  row: H  column: 24
            Class: BAC ends
            High quality sequence stop: 289.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 801 Col=24 Row=H"
                     /sex="M"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

  Query Match             2.8%;  Score 206.8;  DB 1;  Length 289;
  Best Local Similarity   83.3%;  Pred. No. 0;
  Matches  234;  Conservative   25;  Mismatches   18;  Indels    4;  Gaps    3;

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
              :||||||||||||||||||: |||||||| |||||::: |||||| |||||||| |||||
Db        289 AATTATCCAAAGGCACCAGAGCCCTCAGTAAGGAATGTGTCCAGCNTATACTGG-TTATC 231

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAA 3261
              ||||||:|||||||||||||||||||||:|::|||||||||:|||:|||:|||||||  |
Db        230 CTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCATA 171

Qy       3262 WATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACT 3321
              :||||||| |||||||:  ||:||||||||||:|||||:|||||:|||||||||||||||
Db        170 TATGGATT-CCCAGGTACAGCGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACT 112

Qy       3322 CAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCC 3381
              ||||||||||||||||||||:|||||| ||:| ||||||:::|||||||| ||||||||
Db        111 CAGAAAGCCAATACCCATTTAGTAAGATCGACATCTGAAGCAGAAGTGGCTCTCCAGGCC 52

Qy       3382 CCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTG 3422
              |  ||  | | |||  ||||||||:|||||| |||||:|||
Db         51 CTAAAGAAGGCCCT--AACCCAAGCCCCAGTATTAAGCTTG 13


RESULT 365
BF954175
LOCUS       BF954175                 237 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  QV2-NN0045-131100-415-e06 NN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF954175
VERSION     BF954175.1  GI:12371450
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 237)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-NN0045-
            131100-415-e06&t3=2000-11-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 237.
FEATURES             Location/Qualifiers
     source          1. .237
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NN0045"
                     /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

  Query Match             2.6%;  Score 195.8;  DB 1;  Length 237;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  209;  Conservative    1;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       5838 TTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTG 5897
              ||  |||||||||| |||||||||| ||||| |||||||||||||| ||||||||||||
Db          4 TTTCACAGGAAAAATTATTAATCCTAGATGTCCGGGAGGACTTGGAGCCACTGTCTGTTG 63

Qy       5898 GACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGA 5957
              |||||||||||||| ||  |||||||||||||| |||||||||||||||||||| ||||
Db         64 GACTTACTTCACCCATACCAGTATGTCTGATGGGCGTGGAGTTCAAGATCAGGCAGGAGA 123

Qy       5958 AAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTA 6017
              ||||||  |||| ||||| ||||||||||  |||:|||||||| ||||| | | ||||||
Db         124 AAAACACATAAAGGAAGTCATCTCCCAACGGACCCGGGTACATAGCACCCCGAACCCCTA 183

Qy       6018 CAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCC 6070
              |||||||||||||||||||||||||||||||||||||||||| |||||||||| ||
Db         184 CAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTCACC 236


RESULT 366
BF992685/c
```

```
LOCUS       BF992685                195 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  IL5-GN0178-301000-199-f12 GN0178 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF992685
VERSION     BF992685.1  GI:12399008
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 195)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0178-
            301000-199-f12&t3=2000-10-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 193.
FEATURES             Location/Qualifiers
     source          1. .195
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0178"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             2.5%;  Score 187.8;  DB 1;  Length 195;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches  189;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4990 CAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAA 5049
             |||||||||||||||||||| |||||||||||| :||||||||||||||||||||||||||
Db        195 CAGTCCTCAGGGAAGGTCGACAAAATGAATGAATCACTCAAAGGACATCTAAAAAAGCAA 136

Qy       5050 ACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGC 5109
             |||||||||||| |||||||||||||:|||||||||||||||||||||||||||||||||||
Db        135 TCCCAGGAAACCCTCCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGC 76

Qy       5110 AACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATA 5169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 AACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATA 16

Qy       5170 ACCAATGACCTTGTG 5184
             |||||||||||||||
Db          15 ACCAATGACCTTGTG 1


RESULT 367
AQ040405/c
LOCUS       AQ040405                229 bp    DNA     linear   GSS 11-JUL-1998
DEFINITION  CIT-HSP-2328I14.TF CIT-HSP Homo sapiens genomic clone 2328I14,
            genomic survey sequence.
ACCESSION   AQ040405
VERSION     AQ040405.1  GI:3306237
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 229)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2328I14.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .229
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2328I14"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             2.5%;  Score 183.5;  DB 1;  Length 229;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  199;  Conservative   20;  Mismatches   10;  Indels    3;  Gaps    2;

Qy       3184 AGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCR 3243
             || |||||||||||||||||||||:||||||| ||||||||| |||:|::|||||||||:
Db        229 AGTCTATACTGGCTTATCCTCATCCCAAAACCTTAAAGCAACCAAGAGGGTTCCTTGGCA 170

Qy       3244 TAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWAT 3303
             |||:|||:||||||||||:|||||||| |||||||:||||:||||||||||||:|||||:||
Db        169 TAACAGGCTTCTGCCGAATATGGATT-CCCAGGTATGGCGAAATAGCCAGGCCATTATAT 111

Qy       3304 ACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYM 3363
             |||:||||||||||||||||||||||||||||||||:|||||||||:|:||||||::
Db        110 ACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCA 51

Qy       3364 RAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTT 3415
             :|||||||||||||||||  ||  | | |||  ||||||||:|| ||||||
Db          50 GAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCTAGTGTT 1


RESULT 368
BF810306/c
LOCUS       BF810306                235 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  IL5-CI0150-171100-247-g02 CI0150 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF810306
VERSION     BF810306.1  GI:12139295
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 235)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-CI0150-
            171100-247-g02&t3=2000-11-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 234.
FEATURES             Location/Qualifiers
     source          1. .235
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CI0150"
                     /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

  Query Match             2.4%;  Score 177.4;  DB 1;  Length 235;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  199;  Conservative    7;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAAT 4508
             ||||||| ||| || ||||||||||||||| |||||||||||||||||| ||| |||||
Db        235 TCCCCTCAGGATGGCTAGCCACTGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAAT 176

Qy       4509 GGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARA 4568
             |||||||| |||||||||||||| |||||||||||||||||||| |||| |||||||||:|
Db        175 GGAAATTATTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAACACCCATCAGA 116

Qy       4569 TGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGG 4628
             ||||||||||||    |||||||||||||||||||||||||||||||:|||:||||||
Db        115 TGGCCAAATCA----TTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGG 60

Qy       4629 CCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCA 4681
             ||||| ||:|||||||:|:|||||||||||||  :   | | |     ||||
Db         59 CCTGTAAAGTGTGCCAAAGAAATAATCCCCTGCACTGCAGGCCATGCATTTCA 7


RESULT 369
DU639856/c
LOCUS       DU639856                198 bp    DNA     linear   GSS 27-OCT-2005
DEFINITION  Ciuffi-HIV-293T-dKD-49D06.revQ=39 Human Integration Site
            Library-Ciuffi-HIV-293T-dKD Homo sapiens genomic, genomic survey
            sequence.
ACCESSION   DU639856
VERSION     DU639856.1  GI:78202226
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198)
  AUTHORS   Ciuffi,A., Llano,M., Poeschla,E., Hoffmann,C., Leipzig,J.,
            Shinn,P., Ecker,J.R. and Bushman,F.D.
  TITLE     A role for LEDGF/p75 in targeting HIV DNA integration
  JOURNAL   Nat. Med. (2005) In press
COMMENT     Contact: Bushman FD
            Department of Microbiology
            University of Pennsylvania School of Medicine
            402C Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA
            19104-6076, USA
            Tel: 215 573 8732
            Fax: 215 573 4856
            Email: bushman@mail.med.upenn.edu
            The hg17 (May 2004) freeze of the human genome was used.
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .198
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /cell_type="293TdKD (293T stably transfected with 2 shRNAs
                     targeting LEDGF/p75 and HRP2)"
                     /clone_lib="Human Integration Site
                     Library-Ciuffi-HIV-293T-dKD"
                     /note="Sequences cloned using TOPO vectors."

  Query Match             2.3%;  Score 175.2;  DB 1;  Length 198;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  180;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       7284 GGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGG 7343
             ||||||:||| |||||:||||||||| |||||||||||||| ||||||||||||||| ||
Db        198 GGCTGACTAAAAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCCACCTTTAAACATGG 139

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 7403
             ||||||||||||||:|||||||  ||||||||||||||||||||||||||||||||||||
Db        138 GGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 79

Qy       7404 AAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACA 7463
             || |||||||||| |||||||||||| |||:||||||:||||||||||  | |||||||
Db         78 AAAACAGGAGGTAGAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGTAGGGACA 19

Qy       7464 ATGATCGGGATATAAACC 7481
             | ||||||||||||||||
Db         18 AGGATCGGGATATAAACC 1


RESULT 370
DU640205/c
LOCUS       DU640205                198 bp    DNA     linear   GSS 27-OCT-2005
DEFINITION  Ciuffi-HIV-293T-dKD-49D06.fwdQ=42 Human Integration Site
            Library-Ciuffi-HIV-293T-dKD Homo sapiens genomic, genomic survey
            sequence.
ACCESSION   DU640205
VERSION     DU640205.1  GI:78202575
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198)
  AUTHORS   Ciuffi,A., Llano,M., Poeschla,E., Hoffmann,C., Leipzig,J.,
            Shinn,P., Ecker,J.R. and Bushman,F.D.
  TITLE     A role for LEDGF/p75 in targeting HIV DNA integration
  JOURNAL   Nat. Med. (2005) In press
COMMENT     Contact: Bushman FD
            Department of Microbiology
```

```
            University of Pennsylvania School of Medicine
            402C Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA
            19104-6076, USA
            Tel: 215 573 8732
            Fax: 215 573 4856
            Email: bushman@mail.med.upenn.edu
            The hg17 (May 2004) freeze of the human genome was used.
            Class: shotgun.
FEATURES              Location/Qualifiers
     source           1. .198
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /cell_type="293TdKD (293T stably transfected with 2 shRNAs
                      targeting LEDGF/p75 and HRP2)"
                      /clone_lib="Human Integration Site
                      Library-Ciuffi-HIV-293T-dKD"
                      /note="Sequences cloned using TOPO vectors."

  Query Match             2.3%;  Score 175.2;  DB 1;  Length 198;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  180;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       7284 GGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGG 7343
              ||||||:||| |||||:|||||||||| |||||||||||| |||||||||||||||| ||
Db        198 GGCTGACTAAAAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCCACCTTTAAACATGG 139

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 7403
              ||||||||||||||:|||||||  ||||||||||||||||||||||||||||||||||||
Db        138 GGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 79

Qy       7404 AAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACA 7463
              || |||||||||| |||||||||| |||:||||||:|||||||||  | |||||||
Db         78 AAAACAGGAGGTAGAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGTAGGGACA 19

Qy       7464 ATGATCGGGATATAAACC 7481
              | ||||||||||||||||
Db         18 AGGATCGGGATATAAACC 1


RESULT 371
AA077794
LOCUS       AA077794                198 bp    mRNA    linear   EST 24-SEP-1999
DEFINITION  7H01C09 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone
            7H01C09, mRNA sequence.
ACCESSION   AA077794
VERSION     AA077794.1  GI:1837268
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198)
  AUTHORS   Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,
            Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.
  TITLE     2006 expressed-sequence tags derived from human chromosome
            7-enriched cDNA libraries
  JOURNAL   Genome Res. 7 (3), 281-292 (1997)
   PUBMED   9074931
COMMENT     Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
            Tel: 3014020201
            Fax: 3014024735
            Email: egreen@nhgri.nih.gov
            Plate: 01  row: C  column: 09
            Seq primer: -21M13 (ABI).
FEATURES              Location/Qualifiers
     source           1. .198
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="7H01C09"
                      /sex="female"
                      /cell_line="HeLa cell line; ATCC"
                      /lab_host="E. coli strain DH5 alpha"
                      /clone_lib="Chromosome 7 HeLa cDNA Library"
                      /note="Vector: pAMP10; cDNA was generated from cytoplasmic
                      RNA using a mixture of random DNA hexamers and oligo(dT).
                      From this pool of cDNA, human chromosome 7-enriched cDNA
                      was isolated by direct cDNA selection using chromosome 7
                      genomic DNA (cosmids).  The resulting direct-selected cDNA
                      was cloned into a plasmid vector using a non-directional
                      uracil DNA glycosylase (UDG)-mediated cloning strategy."

  Query Match             2.2%;  Score 165.6;  DB 1;  Length 198;
  Best Local Similarity   89.4%;  Pred. No. 0;
  Matches  177;  Conservative    1;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       4855 CATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCT 4914
              |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||||||
Db          1 CATGAAATAATTCCTAGATTTGGACTTCCCCGAGGCTTACAGAGTGACAATGGCCCTGCT 60

Qy       4915 TTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACAC 4974
              ||| ||||  |||||||||||||| |||||||| |||||| |||| ||||||||||||
Db         61 TTCAAGGCTGCAGTAACCCAGGGACTATCCCAGGTGTTAGGCATACAATATCACTTACAC 120

Qy       4975 TGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGA 5034
              ||| |||| |||||||| |||||||| || ||| |||||||||| ||||:|||||||  |
Db         121 TGCCCCTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAACGAAACACTCAAACAA 180

Qy       5035 CATCTAAAAAAGCAAACC 5052
              |||||||||||| ||||
Db         181 CATCTAAAAAAGCTAACC 198


RESULT 372
AQ027014/c
LOCUS       AQ027014                213 bp    DNA     linear   GSS 30-JUN-1998
DEFINITION  CIT-HSP-2313G20.TF CIT-HSP Homo sapiens genomic clone 2313G20,
            genomic survey sequence.
ACCESSION   AQ027014
VERSION     AQ027014.1  GI:3267236
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 213)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2313G20.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .213
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
```

```
                     /clone="2313G20"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             2.1%;  Score 156.5;  DB 1;  Length 213;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  182;  Conservative    2;  Mismatches   29;  Indels    1;  Gaps    1;

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              ||||||||  |||| || ||||||||  |||||||||| |||||| ||||||||||||||
Db        213 CCATTGCCATTCCCGATAGGGCTAAAATCTTGCCATTGTTCCTGCATGGCTAAGTGCCTG 154

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||||| |||||  | ||||||||||  |  | ||||||||| ||||||||||  ||||||
Db        153 GGTTTGTCCTAATCAACCTGAACACTAGTTGCCGGGTTCCACGGTTCTCTTCCATGACCC 94

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              || |||| || ||:| |||||||||| ||  |||||||||||||||||||||||||||||
Db         93 ACAGCTTTTAGTAGAGCTATAACACTCACTGCATGGCCCAAGATTCCATTCCTTGGAATT 34

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACG 481
              |||||||:||| |||| ||||||||||||| |||
Db         33 CGTGAGGCCAA-GAACCCCAGGTCAGAGAACACG 1


RESULT 373
DQ032120
LOCUS       DQ032120                190 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Pan troglodytes HC2312 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ032120
VERSION     DQ032120.1  GI:66883329
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 190)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 190)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .190
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>190
                     /locus_tag="HC2312"

  Query Match             2.0%;  Score 153.5;  DB 1;  Length 190;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  170;  Conservative    5;  Mismatches   15;  Indels    1;  Gaps    1;

Qy        385 CCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGA 444
              |||||||||| |||||:| |||||||||| |||||| ||||||||||||||||||||||
Db          1 CCCACGGCTTGTAATAGAGCTATAACACTCACCACAGGGCCCAAGATTCCATTCCTTGGA 60

Qy        445 ATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGC 504
              ||||||||||:||| |||| |||||||||||||| || |:||||||||||||||||||||
Db         61 ATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCATCTTGGAAGT 119

Qy        505 GGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACC 564
              |||||||| ||:| |||||||||| | ||||||||||||||||||| ||||||||||
Db        120 GGCCTGCTGCCATTTTGGAAGTGGCCCGCCACCATCTTGGGAGCTCTGGGAGCAAGGACC 179

Qy        565 CCCCGGTRACA 575
              ||| |||:|||
Db        180 CCCAGGTAACA 190


RESULT 374
DQ032119
LOCUS       DQ032119                190 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens HC2312 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ032119
VERSION     DQ032119.1  GI:66883328
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 190)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 190)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .190
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
     gene            <1. .>190
                     /locus_tag="HC2312"

  Query Match             2.0%;  Score 150.3;  DB 1;  Length 190;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  168;  Conservative    5;  Mismatches   17;  Indels    1;  Gaps    1;

Qy        385 CCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGA 444
              |||||||||| |||||:| |||||||||| |||||| ||||||||||||||||||||||
Db          1 CCCACGGCTTGTAATAGAGCTATAACACTCACCACAGGGCCCAAGATTCCATTCCTTGGA 60

Qy        445 ATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGC 504
              ||||||||||:||| |||| |||||||||||||| ||||:||||||||||||||||||||
Db         61 ATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGT 119

Qy        505 GGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACC 564
              |||||||  ||:| ||||||||||     |||||||||||||||||||  ||||||||||
Db        120 GGCCTGCCGCCATTTTGGAAGTGGCCTGCCACCATCTTGGGAGCTCTCAGAGCAAGGACC 179

Qy        565 CCCCGGTRACA 575
              ||| |||:|||
```

```
Db        180 CCCAGGTAACA 190


RESULT 375
AI611840/c
LOCUS       AI611840                204 bp    mRNA    linear   EST 15-DEC-1999
DEFINITION  tw16d09.x1 NCI_CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2259857 3',
            mRNA sequence.
ACCESSION   AI611840
VERSION     AI611840.1  GI:4621007
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 958   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 203
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .204
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2259857"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn52"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; This library represents the normalized
                     version of NCI_CGAP_Brn35. Cloned unidirectionally.
                     Primer: Oligo dT. Average insert size 1.19 kb. Tumor
                     types include: meningioma, oligodendroglioma, astrocytoma
                     (grade II), medulloblastoma, astrocytoma (grade IV).
                     Constructed by Life Technologies."

  Query Match             1.9%;  Score 145.9;  DB 1;  Length 204;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  178;  Conservative    3;  Mismatches   14;  Indels   11;  Gaps    2;

Qy       7375 AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATY 7434
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||:
Db        195 AGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAACCATC 136

Qy       7435 TATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGC 7494
             ||||||:|||||||||||  |||||||||||||  |||||||||||| |  ||| |||
Db        135 TATTGCCTGAGAGCACAGTGGGAGGGACAATGAT-TGGATATAAACCCAGGCATTCAAGC 77

Qy       7495 CGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTA 7554
             | ||||||||  |||| ||||||||||||||||||||||||||||||||||
Db         76 CAGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGT---------- 27

Qy       7555 TTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||| |||:|||
Db         26 CTTCACTCTATTAAATCTCGCAACTG 1


RESULT 376
AQ479009/c
LOCUS       AQ479009                183 bp    DNA     linear   GSS 23-APR-1999
DEFINITION  RPCI-11-266H18.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-266H18, genomic survey sequence.
ACCESSION   AQ479009
VERSION     AQ479009.1  GI:4661128
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 183)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-266H18.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .183
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7601945"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-266H18"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"

  Query Match             1.9%;  Score 145.4;  DB 1;  Length 183;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  158;  Conservative    2;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||  ||||| |||||||| |||| |||||||||| |||||| ||| |||
Db        183 TTAGCCCATGCTCCAATGTTGATGACATCGAAGGTACCCCTCCTGGGGAAATATCAACTG 124

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| ||||||| |||| |||||||||||||||||| || ||: | ||||||||||
Db        123 CATGACCCCTACTACACCCCTATTCAGCAGGAAGCAGTTAGCGCAGTCATTGGCCAACCT 64

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |||||||| ||||| |||||||||||||||| |||| |||||||||||||||||||||||
Db         63 CCCCAACAACACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGA 4

Qy       7277 TTT 7279
              |||
Db          3 TTT 1


RESULT 377
BG013341
LOCUS       BG013341                171 bp    mRNA    linear   EST 24-JAN-2001
```

```
DEFINITION  CM4-GN0361-161200-628-a02 GN0361 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG013341
VERSION     BG013341.1  GI:12463434
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM4&t2=CM4-GN0361-
            161200-628-a02&t3=2000-12-16&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 171.
FEATURES             Location/Qualifiers
     source          1. .171
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0361"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             1.9%;  Score 143.7;  DB 1;  Length 171;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  153;  Conservative    3;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5191 CCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTA 5250
              |||||||||||| |||||||||||||||| |||||||||||||||||||||||||||||
Db          2 CCAAGACAGCC-ACTTAGTTGCAGACATTACCTCCTTAGCCAAATATCAACAAGTTCTTA 60

Qy       5251 AAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACA 5310
              ||||||||||||||||||||||||||||||||||||||||||||||||||::||||||
Db         61 AAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACA 120

Qy       5311 TGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCC 5348
              |||||||||||||||||||:||||||||||||||||
Db        121 TGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCC 158


RESULT 378
B57639
LOCUS       B57639                 173 bp    DNA     linear   GSS 20-JUN-1998
DEFINITION  CIT-HSP-2010E10.TR CIT-HSP Homo sapiens genomic clone 2010E10,
            genomic survey sequence.
ACCESSION   B57639
VERSION     B57639.1  GI:2611973
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 173)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .173
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7041132"
                     /db_xref="taxon:9606"
                     /clone="2010E10"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"

  Query Match             1.9%;  Score 141.4;  DB 1;  Length 173;
  Best Local Similarity   82.1%;  Pred. No. 0;
  Matches  142;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
              ||||||||::||||::||||||:||||||||:|||||||| ||||||| |||||: :||||
Db          1 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTATTGGCAAACGGTTGGACTCAT 60

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
              ||||||:||||| || ||||| || ||||::|| |||||||| || ||||||| |||||
Db         61 TGTTTACGGGTAATGATGGCAATATCAGTCTTACTATCTGAAACACTTAAAATCATACAT 120

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              ||:| ||||| |||||||||||||||||||| |:|| ||::||||||||||||
Db        121 GGAATAGATCCTACTGTGTGGACATCTCATTATGTCAATGGCATACTCACTGC 173


RESULT 379
CB298587/c
LOCUS       CB298587                 140 bp    mRNA    linear   EST 28-FEB-2003
DEFINITION  220014_rev_1_A02_r_005.ab1 Chimpanzee testis library Koos Pan
            troglodytes cDNA clone 220014_rev_1_A02_r_005.ab1 5', mRNA
            sequence.
ACCESSION   CB298587
VERSION     CB298587.1  GI:28624017
KEYWORDS    EST.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 140)
  AUTHORS   Hellmann,I., Zollner,S., Enard,W., Ebersberger,I., Nickel,B. and
            Paabo,S.
```

```
  TITLE     Selection on human genes as revealed by comparisons to chimpanzee
            cDNA
  JOURNAL   Genome Res. (2003) In press
COMMENT     Contact: Paabo S
            Evolutionary Genetics
            Max-Planck-Institute for evolutionary Anthropology
            Deutscher Platz 6, 04103 Leipzig, Germany
            Tel: +49-(0)-341-3550 500
            Fax: +49-(0)-341-3550 555
            Email: paabo@eva.mpg.de
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1. .140
                     /organism="Pan troglodytes"
                     /mol_type="mRNA"
                     /db_xref="taxon:9598"
                     /clone="220014_rev_1_A02_r_005.ab1"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="Epicurian Coli (TM) XL-10-Gold"
                     /clone_lib="Chimpanzee testis library Koos"
                     /note="Vector: pUChi; Site_1: SfiI-A; Site_2: SfiI-B; The
                     library was prepared using the SMART cDNA library
                     construction Kit (Clontech), doing only primer extension,
                     but not PCR amplification of the cDNA. The only deviation
                     from the published protocoll was that we cloned the cDNA
                     into a plasmid Vector."

  Query Match             1.8%;  Score 131.8;  DB 1;  Length 140;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        205 CGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARG 264
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        140 CGCCGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGG 81

Qy        265 CGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGC 324
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         80 CGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGC 21

Qy        325 CTGGGTTTGTTCTAATTGAG 344
              | |||||  | |||||||||
Db         20 CCGGGTTNNTCCTAATTGAG 1


RESULT 380
T39519
LOCUS       T39519                   128 bp    mRNA    linear   EST 27-JAN-1995
DEFINITION  ya07b04.r1 Stratagene placenta (#937225) Homo sapiens cDNA clone
            IMAGE:60751 5', mRNA sequence.
ACCESSION   T39519
VERSION     T39519.1  GI:647238
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 128)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  PUBMED    8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: M13RP1.
FEATURES             Location/Qualifiers
     source          1. .128
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:488560"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:60751"
                     /sex="male"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene placenta (#937225)"
                     /note="Organ: placenta; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP
                     XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

  Query Match             1.7%;  Score 127.6;  DB 1;  Length 128;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches  127;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6165 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 6224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 60

Qy       6225 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCT 6284
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db         61 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCT 120

Qy       6285 CACCTGTG 6292
              ||||||||
Db        121 CACCTGTG 128


RESULT 381
CR958717/c
LOCUS       CR958717                 146 bp    DNA     linear   GSS 07-JUN-2005
DEFINITION  Homo sapiens PAC end sequence of RZPDB737C0818D from genomic
            library (orig. Pieter J. de Jong library RPCI-3), genomic survey
            sequence.
ACCESSION   CR958717
VERSION     CR958717.1  GI:66956193
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 146)
  AUTHORS   Schwarz,F., Neubert,P., Schneider,D., Peters,M. and Korn,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2005) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH, Im Neuenheimer Feld 515, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDB737C0818D; RP3-500L14;
            derived from Pieter J. de Jong library RPCI-3;
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDB737C0818D
            RZPDLIB; (Human Genomic Set - RZPD 1.0) RZPD LIB No.737
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=737
            http://www.rzpd.de/products/genomicset/
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
```

```
            Contact RZPD (customer.service@rzpd.de) for further information.
            Clone distribution: http://www.rzpd.de/products/genomicset/
            Seq-primer: T7 (TAA-TAC-GAC-TCA-CTA-TAG-GG)
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .146
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDB737C0818D"
                     /sex="Male"
                     /clone_lib="RPCI-3"
                     /note="Vector: pCYPAC2; RPCI-3 Human Male PAC Library"

  Query Match             1.7%;  Score 123.6;  DB 1;  Length 146;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  132;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1834 GAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACA 1893
              |||| ||||||  ||| ||||||||||||||||||||||||||| ||||||||||||
Db        146 GAAACGCTGGGTGAATCGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACC 87

Qy       1894 CTTTAAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCA 1953
              |||||||||||||||||  | ||||| ||||  ||||| ||||||||||||||||| |||
Db         86 CTTTAAAAAAAGATTGTCCCAATAGAAATAAGTTGCCCCCTCGTCCATGCCCCTTATGTCA 27

Qy       1954 AGGGAATCACTGGAAGGCCCACTGCC 1979
              ||| |||||||||||||||||||||
Db         26 AGGAAATCACTGGAAGGCCCACTGCC 1


RESULT 382
AG197813
LOCUS       AG197813                 118 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-078D17.TJ, genomic survey
            sequence.
ACCESSION   AG197813
VERSION     AG197813.1  GI:45229989
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .118
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-078D17.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"

  Query Match             1.3%;  Score 100.2;  DB 1;  Length 118;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches   99;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229
              |||||||||:|||||||||||||||||||:|||||||:||||:||||||||||||:||
Db          1 CACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAAAG 60

Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAA 4285
              ||||||:|| |||||||||: ||||||||| |||||||||: ||   ||| |:||
Db         61 AGGTGGCAGCCTTACACTGCCTGGGTCATCAGAAAGGAAAGGGAAATAAAAGAGAA 116


RESULT 383
BQ368374/c
LOCUS       BQ368374                 107 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  PM3-GN0501-160501-003-f11 GN0501 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ368374
VERSION     BQ368374.1  GI:21043888
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM3&t2=PM3-GN0501-
            160501-003-f11&t3=2001-05-16&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 107.
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0501"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."

  Query Match             1.0%;  Score 77;  DB 1;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches   77;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       6516 GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGC 6575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGC 48

Qy       6576 ACTAGGTACTGGCATTG 6592
             |||||||||||||||||
Db         47 ACTAGGTACTGGCATTG 31
```

Search completed: August 7, 2006, 13:20:32
Job time : 686 secs

This Page Blank (uspto)

# SCORE OVER LENGTH SEARCHES

Attached is a score over length search. This search was developed to overcome limitations in most standard search systems which favor large sequences with high scoring, but lesser overall identity over smaller sequences with higher overall identity. This search is especially useful for relatively small nucleic acid or polypeptide target sequences (antisense, fragments, probes, primers, RNAi, epitopes, haptens, etc.) claimed functionally via a form of hybridization and/or identity language and having defined upper and lower polynucleotide and or polypeptide length limits.

The score over length search is performed by first running the query sequence using examiner-specified identity and polynucleotide or protein length limit parameters, and saving 65,000 hits and 0 alignments from each desired database. The resulting output is reformatted using a Microsoft Word macro and is imported into Excel. The summary table data are then sorted by the ratio of score of each hit sequence divided by its length and the accession numbers for all hits below the examiner's desired score over length parameters are deleted. The remaining accession numbers are used to pull the corresponding sequences from the databases into subdatabases enriched for good hits and the query sequence is re-run against these subdatabases to yield the final results.

The score over length cutoff for this search is 80%, length 30 - 2527 aa

Examiner Please Note: This cover sheet should be included when submitting results to be scanned.

No suitable results were found in PIR. No results from PIR included here.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2006, 16:20:32 ; Search time 474 Seconds
(without alignments)
2.205 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 7489.6
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 86 seqs, 68934 residues

Total number of hits satisfying chosen parameters: 172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 86 summaries

Issued_Patents_NA

Database : issdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 2889.1 | 38.6 | 2946 | 1 | US-09-175-928-3 | Sequence 3, Appli |
| | 2 | 2551.4 | 34.1 | 2599 | 1 | US-09-719-554-1 | Sequence 1, Appli |
| | 3 | 2315.4 | 30.9 | 2784 | 1 | US-09-719-554-4 | Sequence 4, Appli |
| | 4 | 2204.5 | 29.4 | 2763 | 1 | US-09-949-016-5640 | Sequence 5640, Ap |
| | 5 | 2032.5 | 27.1 | 2055 | 1 | US-09-719-554-22 | Sequence 22, Appl |
| | 6 | 1851.9 | 24.7 | 2304 | 1 | US-08-979-847B-87 | Sequence 87, Appl |
| c | 7 | 1487.1 | 19.9 | 1799 | 1 | US-09-719-554-5 | Sequence 5, Appli |
| | 8 | 1324.2 | 17.7 | 1481 | 1 | US-08-979-847B-105 | Sequence 105, App |
| | 9 | 1320.4 | 17.6 | 1600 | 1 | US-08-979-847B-207 | Sequence 207, App |
| | 10 | 1309.8 | 17.5 | 1600 | 1 | US-08-979-847B-206 | Sequence 206, App |
| | 11 | 1291.9 | 17.2 | 1597 | 1 | US-08-979-847B-205 | Sequence 205, App |
| | 12 | 1201.4 | 16.0 | 1326 | 1 | US-09-719-554-2 | Sequence 2, Appli |
| | 13 | 1117.7 | 14.9 | 1511 | 1 | US-08-979-847B-117 | Sequence 117, App |
| | 14 | 960.2 | 12.8 | 1080 | 1 | US-09-719-554-28 | Sequence 28, Appl |
| | 15 | 952.8 | 12.7 | 1167 | 1 | US-08-691-563C-61 | Sequence 61, Appl |
| | 16 | 952.8 | 12.7 | 1167 | 1 | US-09-374-766-61 | Sequence 61, Appl |
| | 17 | 952.8 | 12.7 | 1167 | 1 | US-08-979-847B-57 | Sequence 57, Appl |
| | 18 | 936.9 | 12.5 | 1158 | 1 | US-08-471-724-1 | Sequence 1, Appli |
| | 19 | 936.9 | 12.5 | 1158 | 1 | US-08-471-969-1 | Sequence 1, Appli |
| | 20 | 936.9 | 12.5 | 1158 | 1 | US-08-384-137-1 | Sequence 1, Appli |
| | 21 | 936.9 | 12.5 | 1158 | 1 | US-08-470-006A-1 | Sequence 1, Appli |
| | 22 | 936.9 | 12.5 | 1158 | 1 | US-08-691-563C-1 | Sequence 1, Appli |
| | 23 | 936.9 | 12.5 | 1158 | 1 | US-09-200-990-1 | Sequence 1, Appli |
| | 24 | 936.9 | 12.5 | 1158 | 1 | US-09-133-411-1 | Sequence 1, Appli |
| | 25 | 936.9 | 12.5 | 1158 | 1 | US-09-374-766-1 | Sequence 1, Appli |
| | 26 | 936.9 | 12.5 | 1158 | 1 | US-08-979-847B-1 | Sequence 1, Appli |
| c | 27 | 894.401 | 11.9 | 1216 | 1 | US-09-719-554-7 | Sequence 7, Appli |
| | 28 | 586 | 7.8 | 741 | 1 | US-08-471-724-9 | Sequence 9, Appli |
| | 29 | 586 | 7.8 | 741 | 1 | US-08-471-969-9 | Sequence 9, Appli |
| | 30 | 586 | 7.8 | 741 | 1 | US-08-384-137-9 | Sequence 9, Appli |
| | 31 | 586 | 7.8 | 741 | 1 | US-08-470-006A-9 | Sequence 9, Appli |
| | 32 | 586 | 7.8 | 741 | 1 | US-08-691-563C-9 | Sequence 9, Appli |
| | 33 | 586 | 7.8 | 741 | 1 | US-09-200-990-9 | Sequence 9, Appli |
| | 34 | 586 | 7.8 | 741 | 1 | US-09-133-411-9 | Sequence 9, Appli |
| | 35 | 586 | 7.8 | 741 | 1 | US-09-374-766-9 | Sequence 9, Appli |
| | 36 | 586 | 7.8 | 741 | 1 | US-08-979-847B-9 | Sequence 9, Appli |
| | 37 | 564.8 | 7.5 | 758 | 1 | US-08-979-847B-112 | Sequence 112, App |
| | 38 | 523.9 | 7.0 | 645 | 1 | US-08-471-724-8 | Sequence 8, Appli |
| | 39 | 523.9 | 7.0 | 645 | 1 | US-08-471-969-8 | Sequence 8, Appli |
| | 40 | 523.9 | 7.0 | 645 | 1 | US-08-384-137-8 | Sequence 8, Appli |
| | 41 | 523.9 | 7.0 | 645 | 1 | US-08-470-006A-8 | Sequence 8, Appli |
| | 42 | 523.9 | 7.0 | 645 | 1 | US-08-691-563C-8 | Sequence 8, Appli |
| | 43 | 523.9 | 7.0 | 645 | 1 | US-09-200-990-8 | Sequence 8, Appli |
| | 44 | 523.9 | 7.0 | 645 | 1 | US-09-133-411-8 | Sequence 8, Appli |
| | 45 | 523.9 | 7.0 | 645 | 1 | US-09-374-766-8 | Sequence 8, Appli |
| | 46 | 523.9 | 7.0 | 645 | 1 | US-08-979-847B-8 | Sequence 8, Appli |
| | 47 | 517.4 | 6.9 | 635 | 1 | US-08-979-847B-102 | Sequence 102, App |
| | 48 | 510.3 | 6.8 | 542 | 1 | US-08-686-878A-48 | Sequence 48, Appl |
| | 49 | 510.3 | 6.8 | 542 | 1 | US-08-721-489-2 | Sequence 2, Appli |
| c | 50 | 467.1 | 6.2 | 601 | 1 | US-09-949-016-44302 | Sequence 44302, A |
| c | 51 | 457.3 | 6.1 | 601 | 1 | US-09-949-016-44303 | Sequence 44303, A |
| | 52 | 436.8 | 5.8 | 443 | 1 | US-09-513-999C-21394 | Sequence 21394, A |
| | 53 | 382.5 | 5.1 | 438 | 1 | US-08-979-847B-201 | Sequence 201, App |
| | 54 | 381.3 | 5.1 | 438 | 1 | US-08-979-847B-128 | Sequence 128, App |
| | 55 | 381.3 | 5.1 | 438 | 1 | US-08-979-847B-129 | Sequence 129, App |
| | 56 | 381.3 | 5.1 | 438 | 1 | US-08-979-847B-131 | Sequence 131, App |
| | 57 | 379.7 | 5.1 | 438 | 1 | US-08-979-847B-130 | Sequence 130, App |
| | 58 | 345 | 4.6 | 429 | 1 | US-08-979-847B-138 | Sequence 138, App |
| | 59 | 344.6 | 4.6 | 429 | 1 | US-08-979-847B-135 | Sequence 135, App |
| | 60 | 341 | 4.6 | 429 | 1 | US-08-979-847B-136 | Sequence 136, App |
| c | 61 | 339 | 4.5 | 410 | 1 | US-09-880-107-538 | Sequence 538, App |
| | 62 | 339 | 4.5 | 429 | 1 | US-08-979-847B-137 | Sequence 137, App |
| | 63 | 306.3 | 4.1 | 374 | 1 | US-08-686-878A-47 | Sequence 47, Appl |
| | 64 | 306.3 | 4.1 | 374 | 1 | US-08-721-489-1 | Sequence 1, Appli |
| | 65 | 238 | 3.2 | 279 | 1 | US-08-686-878A-50 | Sequence 50, Appl |
| | 66 | 238 | 3.2 | 279 | 1 | US-08-721-489-4 | Sequence 4, Appli |
| | 67 | 237.1 | 3.2 | 310 | 1 | US-08-979-847B-100 | Sequence 100, App |
| | 68 | 215.2 | 2.9 | 297 | 1 | US-08-471-724-2 | Sequence 2, Appli |
| | 69 | 215.2 | 2.9 | 297 | 1 | US-08-471-969-2 | Sequence 2, Appli |
| | 70 | 215.2 | 2.9 | 297 | 1 | US-08-384-137-2 | Sequence 2, Appli |
| | 71 | 215.2 | 2.9 | 297 | 1 | US-08-470-006A-2 | Sequence 2, Appli |
| | 72 | 215.2 | 2.9 | 297 | 1 | US-08-691-563C-2 | Sequence 2, Appli |
| | 73 | 215.2 | 2.9 | 297 | 1 | US-09-200-990-2 | Sequence 2, Appli |
| | 74 | 215.2 | 2.9 | 297 | 1 | US-09-133-411-2 | Sequence 2, Appli |
| | 75 | 215.2 | 2.9 | 297 | 1 | US-09-374-766-2 | Sequence 2, Appli |
| | 76 | 215.2 | 2.9 | 297 | 1 | US-08-979-847B-2 | Sequence 2, Appli |
| | 77 | 139.2 | 1.9 | 150 | 1 | US-08-979-847B-36 | Sequence 36, Appl |
| | 78 | 79.2 | 1.1 | 111 | 1 | US-08-471-724-7 | Sequence 7, Appli |
| | 79 | 79.2 | 1.1 | 111 | 1 | US-08-471-969-7 | Sequence 7, Appli |
| | 80 | 79.2 | 1.1 | 111 | 1 | US-08-384-137-7 | Sequence 7, Appli |
| | 81 | 79.2 | 1.1 | 111 | 1 | US-08-470-006A-7 | Sequence 7, Appli |
| | 82 | 79.2 | 1.1 | 111 | 1 | US-08-691-563C-7 | Sequence 7, Appli |
| | 83 | 79.2 | 1.1 | 111 | 1 | US-09-200-990-7 | Sequence 7, Appli |
| | 84 | 79.2 | 1.1 | 111 | 1 | US-09-133-411-7 | Sequence 7, Appli |
| | 85 | 79.2 | 1.1 | 111 | 1 | US-09-374-766-7 | Sequence 7, Appli |
| | 86 | 79.2 | 1.1 | 111 | 1 | US-08-979-847B-7 | Sequence 7, Appli |

ALIGNMENTS

RESULT 1
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

```
;  FILE REFERENCE: 6006B.AJ172A
;  CURRENT APPLICATION NUMBER: US/09/175,928A
;  CURRENT FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-175-928-3

  Query Match            38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity  98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              ||:||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
              ||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||
Db        606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905

Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
              |||||||||||||||| ||||||||:|||||||||||||||||||||:||:||||:
Db        906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
              |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
              ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
              |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684

Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
              |||||||||:|||:|||||||||||||||||||||||||| ||||||||||||||||||
Db       1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924

Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984
```

```
Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
             |||||||||||||||||||||||:|||||||||||||||||||||:|||||||||||||:|
Db       2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104

Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164

Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
             |||||||||||||||||:|||||||||:|||||||||||||||||:||||||:|||||||
Db       2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
             ||:|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
             |||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db       2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy       7538 GCTCTGTTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||||||||||||||||||||:||||:
Db       2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929


RESULT 2
US-09-719-554-1
; Sequence 1, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 2599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-719-554-1

  Query Match             34.1%;  Score 2551.4;  DB 1;  Length 2599;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2552;  Conservative   39;  Mismatches    5;  Indels    2;  Gaps    2;

Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714
             |||   | | |||||||||||||||||||:||||||:||||||||||||||||||:|:||
Db          4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy       4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123

Qy       4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
             |||||:|||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183

Qy       4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243

Qy       4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303

Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363

Qy       5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483

Qy       5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543

Qy       5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 603

Qy       5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
             |||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db        604 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 663

Qy       5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
             |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        664 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 723

Qy       5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 783

Qy       5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
```

```
Db        784 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 843

Qy       5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554

Db        844 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 903

Qy       5555 AAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTG 5614

Db        904 AAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTG 962

Qy       5615 TTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 5674

Db        963 TTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 1022

Qy       5675 CCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 5734

Db       1023 CCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 1082

Qy       5735 ATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 5794

Db       1083 ATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 1142

Qy       5795 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 5854

Db       1143 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 1202

Qy       5855 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 5914

Db       1203 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 1262

Qy       5915 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 5974

Db       1263 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 1322

Qy       5975 TAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 6034

Db       1323 TAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 1382

Qy       6035 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 6094

Db       1383 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 1442

Qy       6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 6154

Db       1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 1502

Qy       6155 TGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 6214

Db       1503 TGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 1562

Qy       6215 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATA 6274

Db       1563 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATA 1622

Qy       6275 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 6334

Db       1623 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 1682

Qy       6335 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 6394

Db       1683 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 1742

Qy       6395 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 6454

Db       1743 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 1802

Qy       6455 TCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 6514

Db       1803 TCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 1862

Qy       6515 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 6574

Db       1863 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 1922

Qy       6575 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 6634

Db       1923 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 1982

Qy       6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694

Db       1983 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 2042

Qy       6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754

Db       2043 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 2102

Qy       6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814

Db       2103 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 2162

Qy       6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874

Db       2163 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 2222

Qy       6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934

Db       2223 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 2282

Qy       6935 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCC 6994

Db       2283 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCC 2342

Qy       6995 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGC 7054

Db       2343 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGC 2402

Qy       7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATG 7114

Db       2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATG 2462

Qy       7115 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 7174

Db       2463 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 2522

Qy       7175 CCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGT 7234

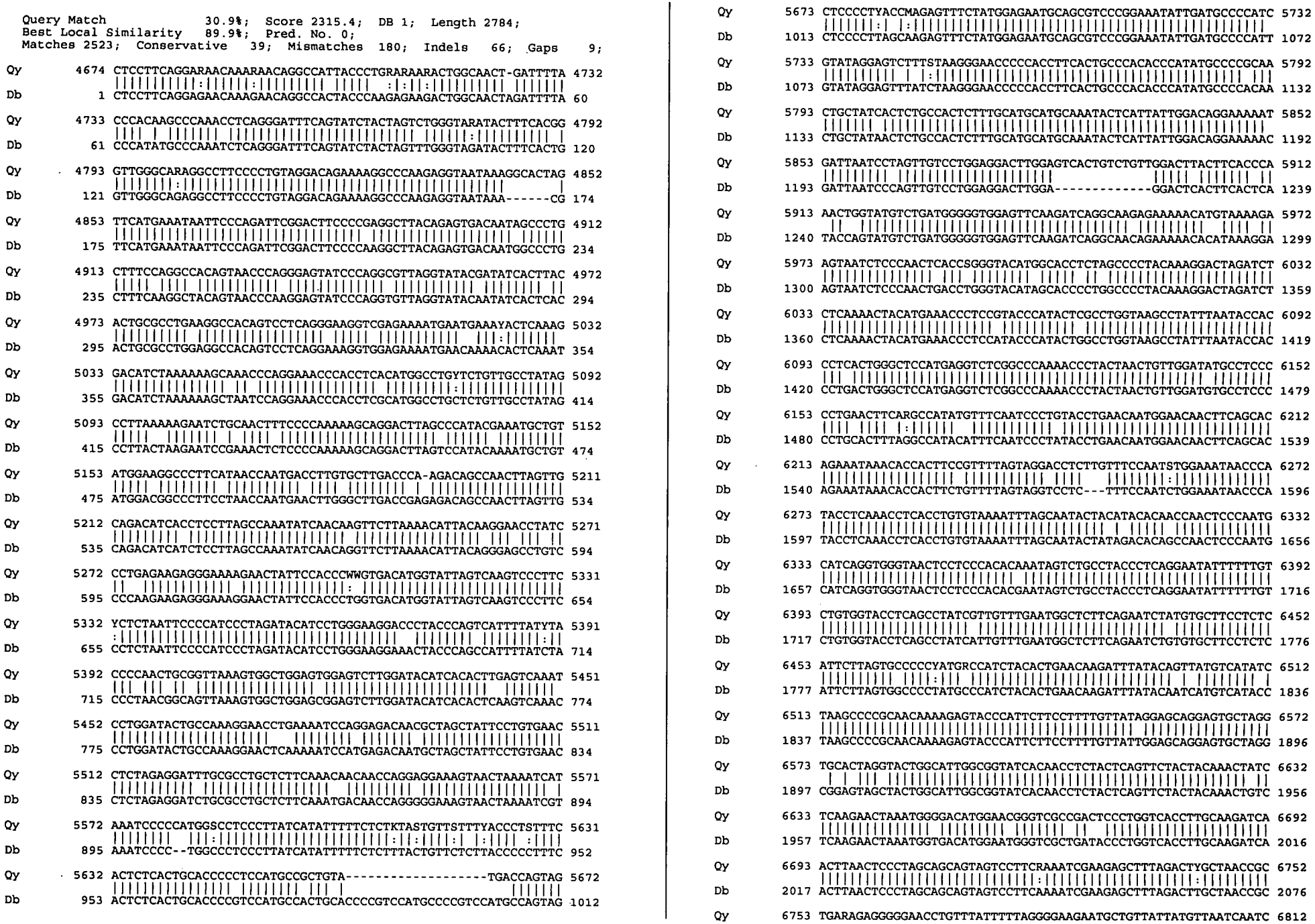
Db       2523 CCAATTCAGCAGGAAGCAGTTAGAGCGGTC-TCGGCCAACCTCCCCAACAGCACTTAGGT 2581

Qy       7235 TTTCCTGTTGAGATGGGG 7252

Db       2582 TTTCCTGTTGAGATGGGG 2599
```

RESULT 3
US-09-719-554-4
; Sequence 4, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUM
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2784
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-4

Query Match 30.9%; Score 2315.4; DB 1; Length 2784;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 2523; Conservative 39; Mismatches 180; Indels 66; Gaps 9;



```
Db       2077 GGAAAGCGGGGGAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATC 2136

Qy       6813 CGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGA 6872

Db       2137 CGGAATCATCACCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGA 2196

Qy       6873 GCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTT 6932

Db       2197 GCTGCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTT 2256

Qy       6933 CTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCT 6992

Db       2257 CTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCT 2316

Qy       6993 CCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAAT 7040

Db       2317 CCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAAT 2376

Qy       7041 GGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAG 7100

Db       2377 GGAGCCCCAGATGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAG 2436

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160

Db       2437 CCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACA 2496

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220

Db       2497 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC 2556

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280

Db       2557 AACAGCAGTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTC 2616

Qy       7281 CTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACA 7340

Db       2617 CTAGACCAACTAAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACA 2676

Qy       7341 CGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAAT 7391

Db       2677 CCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAAT 2736

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTG 7439

Db       2737 GCTAATTAGGCAAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTG 2784
```

RESULT 4
US-09-949-016-5640
; Sequence 5640, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5640
; LENGTH: 2763
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5640

Query Match 29.4%; Score 2204.5; DB 1; Length 2763;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2194; Conservative 35; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411

Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471

Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531

Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591

Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651

Db        774 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711

Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771

Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831

Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891

Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951

Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011

Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071

Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131

Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191

Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251

Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311

Db       1434 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371

Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
```

```
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491

Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551

Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611

Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671

Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731

Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791

Db       1914 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851

Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911

Db       2034 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971

Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031

Db       2154 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091

Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151

Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211

Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271

Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331

Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391

Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451

Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511

Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571

Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGC 7581

Db       2754 TTGCAACTGC 2763
```

```
RESULT 5
US-09-719-554-22
; Sequence 22, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 2055
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2055)
;   OTHER INFORMATION:
US-09-719-554-22

  Query Match             27.1%;  Score 2032.5;  DB 1;  Length 2055;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 2025;  Conservative   30;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249

Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309

Db         61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369

Db        121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429

Db        181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489

Db        241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549

Db        301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609

Db        361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669

Db        420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479
```

```
Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
             ||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
             ||||||||||||||||||||||||||||||:||||||||||||||||||:||||||||||
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
             ||||||||||||||||||||||||||:|||||||||:||||||||||||||||:||||||
Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
             :||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245
             ||||||||||||||||
Db       2040 TAGGTTTTCCTGTTGA 2055
```

```
RESULT 6
US-08-979-847B-87
; Sequence 87, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
```

; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2304 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 87:
US-08-979-847B-87

Query Match 24.7%; Score 1851.9; DB 1; Length 2304;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2032; Conservative 119; Mismatches 150; Indels 29; Gaps 15;

```
Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             |||||||||||||||||||||||||| |||||||| |||| ||||||||||||||| |||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| ||||||| |||||||||||| |||||||||  |   |||||||||||||||| |
Db         59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             | | ||||| |||||||| | ||||||| |||| |||||||||||||||||||||| |
Db        119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                | |||  | || ||| | ||||||||| |||| ||:|| |||||||||||||| || |
Db        179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
             ||||||| ||||| |||||||||||||||||||||||||||| |||||||||||||| ||||
Db        234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db        294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
             || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
             |||||||||||||||| || |||||||||||||||||||||||||||||||||||||||||
Db        413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
                | |||||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db        472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
             |||||||||||||||| |||||||||||||||| |||||||||||||||  |  ||||||| |
Db        532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
             | | |         |||   | | |  |||||||| |||||||||| ||||||||||||
Db        592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
             |||||||||||| ||| ||||| ||| | | | |||||||||||||| ||||||||||||
Db        644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
             ||| ||::||||| |||||||||||||||||||||||||||||:||||||||||:|||||
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||| ||
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
             |  ||||||:| |||||||||||||||||||:|||||||||||||||||||||||||:|||
Db         823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
             ||||||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db         883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
             |||||||:||| |||||||||||||||||:| :|||::|||||||||||||||||||||||
Db         943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
             |||||||||| :|||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:
Db        1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
             |||| :|||||||||||| || ||| |||||||||:||||:|||||||||||||||||:|
Db        1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
             |||||||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||
Db        1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
             ||||||||:|::|||||||||:|||:|||:||||||||||:|  |||| |||||||:
Db        1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
             |  |||||||||  :|||||:|4|||:|||||| |||||||||||||||||||||||| ||||
Db        1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
             |:|||||||||:|:||     :::|||||||||||||||||||  ||  | | |||  ||||
Db        1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
             ||||:||||||||| ||:||||||||:|||||||| ||||:||:|||:: ||||| ||||
Db        1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
             |||||:||:|||||||:|||||||||||| |||:|| |||||:|||||||||||::|||
Db        1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
             ::||||||:|||||||:|||||||||||||||||||||||:|:|||||||||||||||| |
Db        1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
             || |||||||||||::|||||||||||||||||||||||||||||||||:|||||||||||
Db        1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
             ||||||||||||||||||:||||::|||||||||||||||||||||||||||||| ||||
Db        1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
             |||||||: |||||||||  |:|||||||||||||||||||: |||||||||| :|||| |
Db        1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
             |||||||||||||||||| || |:| |||||||||||||||||||||||||||||:||:|
Db        1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
             ||||||||||||:||||| |||||||||||||  |||||||||||||||:||||:|||||
Db        1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
             :|||||||||:|||| |||||||||||||||||||||||||||||| | ||||||||||
Db        1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894
```

```
Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
             |||||||||:|||||||||||||| ||||||||||||||||||:|||||||||||| |||||
Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
             |||||||||||||||||:| ||:||||||||||||||:|||:|:|||||||||||||||
Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
             ||||||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| ||||||
Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
             ||||||::|||||||||:|:||| |||||||||||||||||:|||:|||||||||||: |
Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
             ||||||||||||||||||:|||||||:||||:|||||||||||||:|||||||||:|||||
Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: ||||||||| ||| | ||:||||||||||||:|||: ||: ||||||| :
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
             ||||||||||:|||| ||| |||||||||||| ||||||||||||||||
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 7
US-09-719-554-5/c
; Sequence 5, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 1799
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-719-554-5

  Query Match             19.9%;  Score 1487.1;  DB 1;  Length 1799;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches 1613;  Conservative   27;  Mismatches  116;  Indels   45;  Gaps    5;

Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||| ||||| ||||||||:|||||||||||||||||
Db       1799 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 1741

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||||||||||
Db       1740 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 1681

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
               ||||| |||||||||:|||:|||||||||||| ||||||| || |||| ||||||||||
Db       1680 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 1621

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
             ||||||||||  |||||||| |  :||| |||| | ||||||||||||||||||||||||
Db       1620 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 1561

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
             |||||| |||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 1501

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
             |||| ||||||||||||||||||||||||||||||| ||||||||| |||||||||| ||||||
Db       1500 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 1441

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
             |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1440 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 1381

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAG 6022
             ||||||| ||||||| ||||||||| |||:|||||||| ||||| ||||||||||||||||
Db       1380 ATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCTACAAAG 1321

Qy       6023 GACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTAT 6082
             |||||||||||| |||||||||||||||||||| |||||||||| ||||||||||||||||
Db       1320 GACTAGATCTCTTAAAACTACATGAAACCCTCCATACCCATACTTGCCTGGTAAGCCTAT 1261

Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGA 6142
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1260 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAAACCCTACTAACTGTTGGT 1201

Qy       6143 TATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACA 6202
             | |||||||||||| | ||||:|||||   ||||||||||||||||||||||||||||||
Db       1200 TGTGCCTCCCCCTGTATTTCAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACA 1141

Qy       6203 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1140 ACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGG 1081

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
             |||||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db       1080 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCA 1021

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAA 961

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        960 TATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 901

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
             |||||||||||||||||||||||| |||:|||| |||||||||||||||||||||| ||
Db        900 GCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATT 842

Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
             |||| || |||||||| |||||||||||||| |||||||||||||||||| |||||||
Db        841 ATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAG 782

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
             |||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||||||
Db        781 GAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACT 722

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
             |||||||||||||||||| ||||| |||||||| |||| ||||||||||||||||||||
Db        721 ACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCT 662

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
             |||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||||||
Db        661 TGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTAGACT 602

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGG---------------------- 6782
             :||||||| ||||:|||||||||| ||||||||||||||
Db        601 TGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATG 542

Qy       6783 ----------GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAG 6832
                       ||||||||| |||||||||||||||||||||| |||| ||||| |||||||
Db        541 TTATTTTAGCGGAAGAATGTTGTTATTATGTTAATCAATCCTGAATTGTCACAGAGAAAG 482

Qy       6833 TTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC-GAAACACTGGACCC 6891
             ||:||||||||:|||| ||||||||||||:| |||||:||||||  ||||||  |||||
```

```
Db        481 TTGAAGAAATTCGAGATTGAATACAACGTAGAACAGAGGAGCTTCAAAAACACCAGACCC 422

Qy       6892 TGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCT 6951
              |||||||||||||||:|||||||||||||||||||||||||||||| ||||||||||
Db        421 TGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGATCTCTAGCAGCT 362

Qy       6952 ATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCT 7011
              |||||||| |||||||||||||||||||||||:|||||||||||| |||||||||
Db        361 CTAATATTGATACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCT 302

Qy       7012 TCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGATGCAGTCC 7059
              ||||||||| ||| |||:|| |||            |||||||| ||| ||||| |||||
Db        301 TCCAGAATCAAAGTTGTAAAGCTACAAATCGTTCTTCAAATGGAACCCCAGATGAAGTCC 242

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
              | ||||||||||||||  |||||||||||||||| :||||||| | ||  ||  ||||
Db        241 ATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTACTAGCCCATGCTCCAATTGTAAT 182

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
              || ||| || |||||||||| |||||||||| |||||||||| |||||| ||||||||
Db        181 GATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAAT 122

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
              || |||||||||||||||  |||:||| |||||||||||||||||||| |||||||||
Db        121 TCCGCAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCCAACAGCACTTGGGTTTTCC 62

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||  |||||||||||||||||| |||||||||||||||||| ||:|||||||||
Db         61 TGTTGAGTGGGGGGACTGAGAGACAGGATTAGCTGGATTTCCTAGGCCGACTAAGAATCC 2

Qy       7300 Y 7300
              :
Db          1 C 1
```

```
RESULT 8
US-08-979-847B-105
; Sequence 105, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE:*PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 105:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1481 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-08-979-847B-105

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||| |||||| ||||||||||||||:|||:|||||||||||| |||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              | || || ||| |||||||||||||||||||| ||||||||| | |:|||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||| |||||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||| |||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||| |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||| |||||||||||||||||||||||| |||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||||||| |||||||||||| ||||||:|||||  ||||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||| | ||||||||| |||||||||||||||||||||||| |||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840
```

```
Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             |||||||||||||||||| | ||||| || ||||||||| ||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             |||||||||||||  ||||||||||||||    ||||||||||||||||||| |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||| ||||||| |||||||||| |
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || ||||||| |||||||||||||||||||||||| |||| ||||||||||||||||:|||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:||||||:||||||| |||||| | || ||||||||||
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| ||||||||||||||||||||||||||||| ||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              ||||||||||||||||||||||||:||||||||||||| ||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
             ||||||||:|| ||||| ||||      | ||| |  || |
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481


RESULT 9
US-08-979-847B-207
; Sequence 207, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 207:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-08-979-847B-207

  Query Match             17.6%;  Score 1320.4;  DB 1;  Length 1600;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches 1402;  Conservative  111;  Mismatches   87;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| ||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||  | |||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| |||||||   |||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |         |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||||| |||  |||||||:| ||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:||||||:|:|||||| :|||||||||||||||:|:::
```

```
Db        650 GYCARTYYTCATACCTGGACAYTCTTGTYCYTCRGTACRTGGATGATTTACTTTTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
              ||:|||||||||||||||||||||||||||||||||||::|||||:|:||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGMACTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
              |||||||||:|||||||||||:|:|:||||::|||||||||||||:||:|||:||||||||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCRGCTCTGCTCACAGSAGRTTARATACTTAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
              |:||||:|||||||||||||||||:::|||||||||||||||:::|||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGRRCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
              :|||||||||||:|||||||||||||||||||::|::|||||||||:|||:|||:||||
Db        890 STTATCCTCATCCCAAAACCCTAAAGCAACTAASAGSGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
              |||:||||::|||| |||||||:  :|::::|||||||| :||||||||||::||||||:|
Db        950 GCCRAAWAYRGATT-CCCAGGTACASCMMRRTAGCCAGACCATTAWATACACKAATTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
              ||||||||:|||||||:||||:|||||:|||||||||:|:||||||:::|||||||||||
Db       1009 GAAACTCARAAAGCCARTACCYATTTAGTAAGATGGACAYCTGAAGCAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
              |||||||  ||  | | |||  ||||||||:|||||||||  ||:|||||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
              | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
                |||||:|||||||||||::||| ::||||||:|||||||||:||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
              |||||:|:|||| |||||:||||| ||| ||||||||||||:: ||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
              |||||||||||||:|||||||| ||||||||||||||||||||:|||||::|||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
              |||||||||||||||||||| |||||||||||: ||||||||| |:|||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
              ||||||: |||||||||| :|||| |||||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
              |||||||||||||||||||:||:|||||||||||:|||| ||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 10
US-08-979-847B-206
; Sequence 206, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 206:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-08-979-847B-206

  Query Match             17.5%;  Score 1309.8;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
              ||||||| |||||||||||||| ||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
              |||||||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
              |||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
              |||||||||||||||||||||||||  | |||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
              ||||| |||||| |||||||   ||||||||||||||||| ||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
              |||||||||||  |  ||||||| || | |         |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
              |  |||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409
```

```
Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
Db        650 GYCARTYCTCATACCTGGACAYTCTTGTYCYTCAGTAYRKGGATGAYTTAMTTWTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
Db        710 CCCRTTCAGAAACCTTGTGSCAMCAAGCCACCCAAGHRCTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
Db        770 GTGGCTACAAGGTTTCCAAACMAARGGCTCRSCTCTGCTCACASSAGRTTARATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
Db        830 GGCTAAAATTATCCAAAGKCRCCAGGGCCCTCAGWGAGGAACGTATCCAGCSTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
Db        890 MTTATCCMCATCCCAWAACCMTAAAGCAACTAAGARGGTTCCTTGGCATAWCAGSYTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
Db        950 GCCGAAWAYRGATT-CCCVGRTACASYSMAATAGCCAGRCCATTATRTACAYTADYTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
Db        1009 GAAACTCAGAAAGCCAATACCYATWTAGTAAGATGGACACCTGARACAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
Db        1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
Db        1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
Db        1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
Db        1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
Db        1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
Db        1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
Db        1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
Db        1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
Db        1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 11
US-08-979-847B-205
; Sequence 205, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 205:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1597 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 205:
US-08-979-847B-205

  Query Match             17.2%;  Score 1291.9;  DB 1;  Length 1597;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches 1405;  Conservative   88;  Mismatches  104;  Indels   19;  Gaps   10;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
Db          61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
```

```
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535

Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595

Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655

Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715

Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775

Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835

Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895

Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955

Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015

Db        650 GTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCG 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075

Db        710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135

Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195

Db        830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255

Db        890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315

Db        950 GCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375

Db       1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435

Db       1066 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1123

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495

Db       1124 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1182

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555

Db       1183 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1242
```

```
Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615

Db       1243 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1302

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675

Db       1303 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1362

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735

Db       1363 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1422

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795

Db       1423 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1482

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855

Db       1483 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1542

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911

Db       1543 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1597
```

RESULT 12
US-09-719-554-2
; Sequence 2, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-2

Query Match 16.0%; Score 1201.4; DB 1; Length 1326;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1295; Conservative 16; Mismatches 14; Indels 12; Gaps 9;

```
Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777

Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837

Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897

Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957

Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017

Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300
```

```
Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
             |:||||::|||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
             ||||||||||:||||||||||||||||||||||||||||||      ||||||||
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
             |||| |||||||| ||||||||||||||||||||:|||||||||||||||:|||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
             |||||   ||| | |     ||||| ||||||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711

Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
             |||||:|||||||||||||||||:|||| ||:||||:|||||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
             |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009

Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
             |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
             |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326
```

```
RESULT 13
US-08-979-847B-117
; Sequence 117, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 117:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1511 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-08-979-847B-117

  Query Match            14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity  87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| |||||||||
```

```
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | ||||||||||||||| ||::|||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||| |||| |||| ||||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| ||||||||||:|||||||||||||||||||||||| ||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||| |||   |   |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      ||||||||||||:||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:|||||| |||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||| ||||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||||| ||||||||      | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| ||||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||| ||||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

RESULT 14
US-09-719-554-28
; Sequence 28, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1080)
; OTHER INFORMATION:
US-09-719-554-28

Query Match 12.8%; Score 960.2; DB 1; Length 1080;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1052; Conservative 16; Mismatches 11; Indels 12; Gaps 9;

```
Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
              |||||||||||||||||||||||||||||||||||||||||::|||||||||||||||||
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
              ||||||||:||||::|||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
              ||||||:|||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419
```

```
Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
             |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||     |
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
             |||||||||| |||||||| |||||||||||||||||||:|||||||||||||||||||:
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
             |||||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
             ||||||||||||:||||||||||||||||:|||| ||:||||:||||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAAGGCCACCGCTTTA 1789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
             ||||||||||||
Db       1070 TGAATGCCTAA 1080


RESULT 15
US-08-691-563C-61
; Sequence 61, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1167 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-61

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
             |||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::||| |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
             |||||||||| ||  ||  |||  ||||||||:||||||||||||:||||||||:||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::||||||  ||||||||:||:||||||:|||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: |||||||||||::||||::||||||:||||||||:||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             ||||||||||||||||:|||||||||||||||||||||||||||||:|||||::|||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||||  |||||||||||||||||||||||||||: ||||||||||| |||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: ||||||||||||:||||  || |||||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||:||:|||||||||||||||:||||| ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  |||||||||||||:|||||:|||||:|||||||||||:|||| ||||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||||| ||||||:|||||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||| |||||||||||||||||||||:| ||:||||||||
```

```
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092

Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152

Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211

Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271

Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331

Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391

Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448

Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481

Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 16
US-09-374-766-61
; Sequence 61, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1167 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-374-766-61

  Query Match            12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity  86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372

Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432

Db*        61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492

Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552

Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612

Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672

Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732

Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792

Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852

Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912

Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972

Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032

Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092

Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152

Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
```

```
           |||||||||:|||:|||||||||||: ||||| |||||||||||||:|||||||:|||
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:|||||||:|||||||||||: |||||||||  ||| | ||:
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             |||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || |||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             |||||||  || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 17
US-08-979-847B-57
; Sequence 57, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 57:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1167 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 57:
```

US-08-979-847B-57

Query Match 12.7%; Score 952.8; DB 1; Length 1167;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1013; Conservative 77; Mismatches 73; Indels 10; Gaps 6;

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
              |||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::||| |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
              |||||||||  ||  ||  |||  ||||||||:||||||||||||:||||||||:||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
              |||||||||:||:|||::||||||  ||||||||:||:||||||:|||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
              :|||:|||||: ||||||||||::||||::||||||:||||||||:||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
              ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
              |||||||||||||||||:|||||||||||||||||||||||||||||:|||||::|||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
              |||||||||||||  ||||||||||||||||||||||||||: |||||||||| |||| ||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
              |||||||||: |||||||||||:||||   || ||||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
              |||||||||||||||||||||||:||:|||||||||||||||:||||| ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
              ||  |||||||||||||:|||||:|||||:||||||||||||:|||| ||||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
              |||||||||||||| | |||||||||||| |||||:|||||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
              |||||||||:|||||||||||| ||||||||||||||||||||:| ||:||||||||
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
              |||||||:|||:|||||||||||||||||||||:||||||||:|||||||||| ||||
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
              ||||:|:||||||:|||||||| ||||||||||::||||||||:|:||| ||||||||
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
              |||||||||:|||:|||||||||||: ||||| |||||||||||||:|||||||:|||
Db         835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
              |:|||||||||||:|||||||:|||||||||||: |||||||||  ||| | ||:
Db         895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
              ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db         955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014
```

```
Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || ||||||||||||||||||||  ||    ||| |||||| |||||||||| || ||||||
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             |||||||  || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 18
US-08-471-724-1
; Sequence 1, Application US/08471724
; Patent No. 5800980
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,724
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-724-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  |||||||| |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||

Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||||| |||:||  |||||:|||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 19
US-08-471-969-1
```

```
; Sequence 1, Application US/08471969
; Patent No. 5871745
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-969-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:|||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|

Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||||||| |||||:||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:|||||  ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:|||||||| ||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 20
US-08-384-137-1
; Sequence 1, Application US/08384137
; Patent No. 5871996
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
```

```
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22314
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/384,137
;       FILING DATE:  February 6, 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Berridge, William P.
;       REGISTRATION NUMBER:  30,024
;       REFERENCE/DOCKET NUMBER:  WPB 36055
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-836-6400
;       TELEFAX:  703-836-2787
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1158 bases
;       TYPE:  nucleotide
;       STRANDEDNESS:  single-stranded
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-384-137-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db       1150 CTTAAACC 1157


RESULT 21
US-08-470-006A-1
; Sequence 1, Application US/08470006A
; Patent No. 5962217
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND
;    TITLE OF INVENTION:  BIOPOLYLMER CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,006A
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055B
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-470-006A-1

  Query Match            12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| |||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||||:|||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||||| |||:||  |||||:||||||||||::||| ::||||||:||||

Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 22
US-08-691-563C-1
; Sequence 1, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-1

  Query Match            12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;
```

```
Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db       1150 CTTAAACC 1157
```

RESULT 23
US-09-200-990-1
; Sequence 1, Application US/09200990
; Patent No. 6184025
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Francois MALLET
; APPLICANT: Bernard MANDRAND
; APPLICANT: Frederic BEDIN
; APPLICANT: Frederic BESEME
; TITLE OF INVENTION: MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
; TITLE OF INVENTION: INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
; TITLE OF INVENTION: CONSTITUENTS THEREOF (AS AMENDED)
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/200,990
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/471,969
; FILING DATE: June 6, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36055A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1158 bases
; TYPE: nucleotide
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-200-990-1

Query Match 12.5%; Score 936.9; DB 1; Length 1158;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 990; Conservative 80; Mismatches 69; Indels 9; Gaps 6;

```
Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196
```

```
Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             ||||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             |||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:|||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:||  |||||:|||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| |||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 24
US-09-133-411-1
; Sequence 1, Application US/09133411
; Patent No. 6342383
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOLY
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/133,411
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-133-411-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  |||||||| |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             ||||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
```

```
                |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                |||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :|||||:|||||:|||||| |||||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
                | ||:||||||||:||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:|||||||||||| |||:|| |||||:||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                ||||| ||
Db       1150 CTTAAACC 1157


RESULT 25
US-09-374-766-1
; Sequence 1, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-374-766-1

  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  |||||||| |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||||:|||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556
```

```
Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db       1150 CTTAAACC 1157


RESULT 26
US-08-979-847B-1
; Sequence 1, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1158 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-979-847B-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

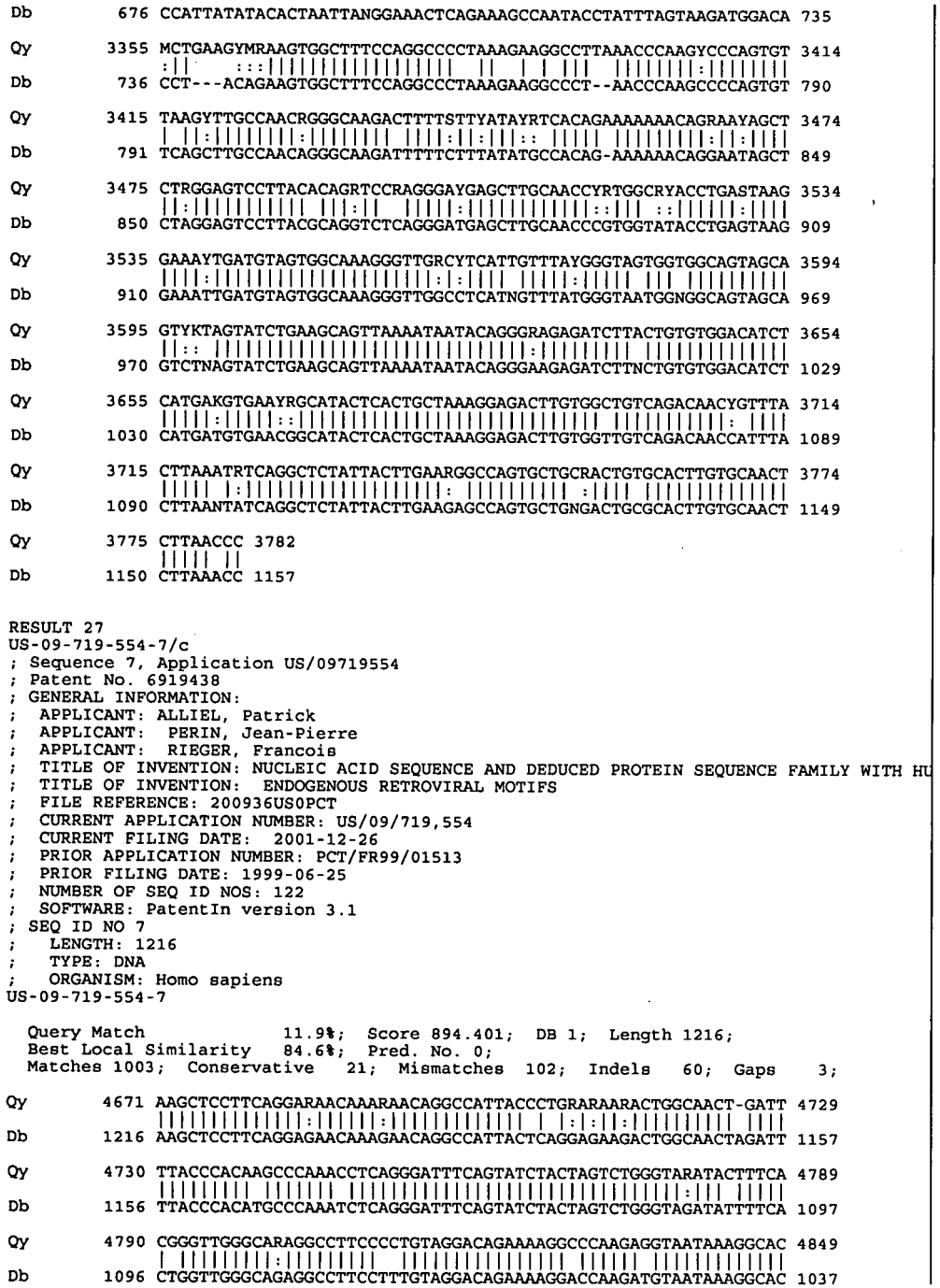
Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

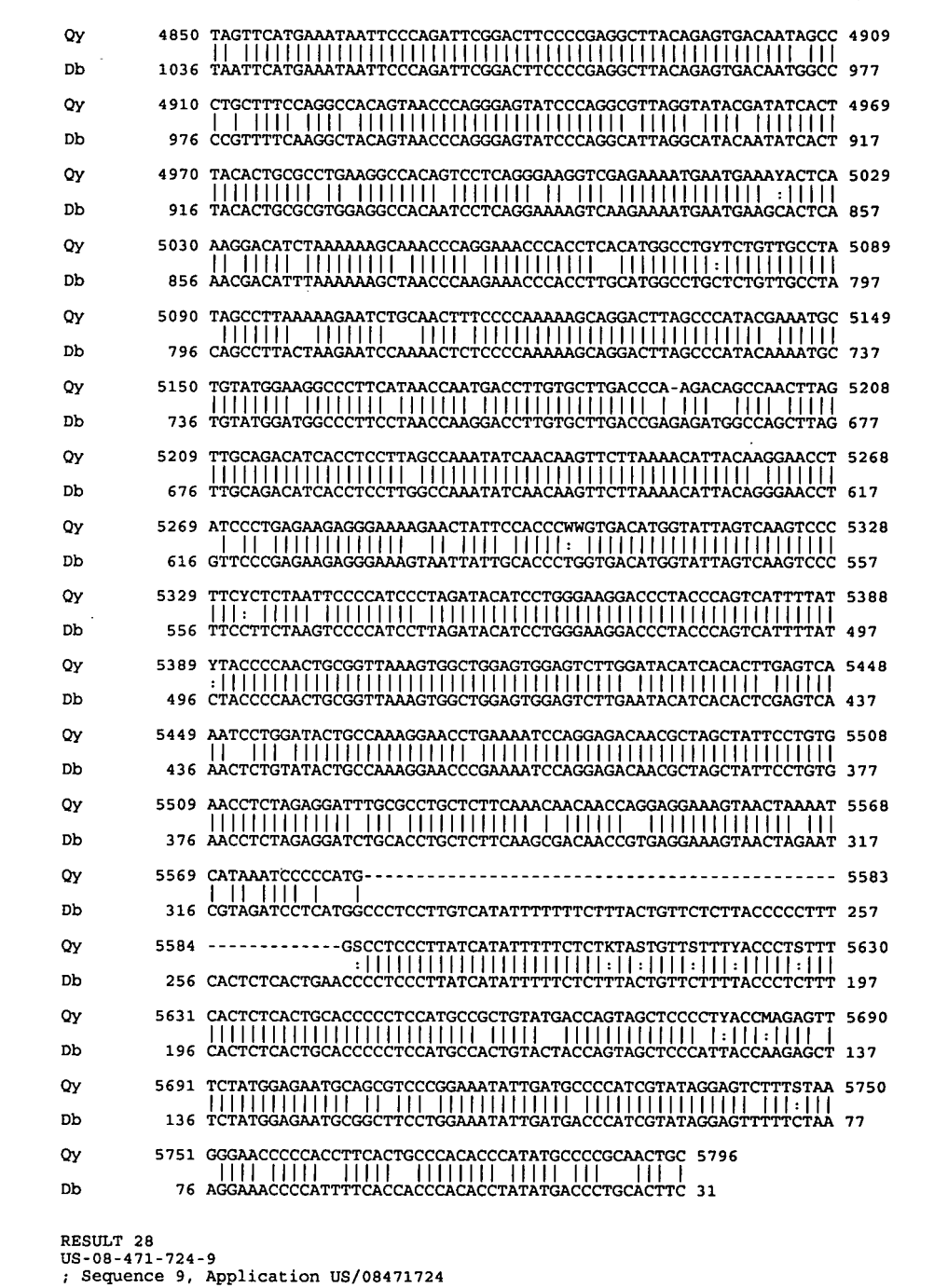
Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
```



RESULT 27
US-09-719-554-7/c
; Sequence 7, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-7

Query Match 11.9%; Score 894.401; DB 1; Length 1216;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1003; Conservative 21; Mismatches 102; Indels 60; Gaps 3;





RESULT 28
US-08-471-724-9
; Sequence 9, Application US/08471724

```
; Patent No. 5800980
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,724
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-724-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||  :|||||:|||||||||  |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:|||||||||:|      | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||||| ||:|||||||:|||||||| ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: ||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 29
US-08-471-969-9
; Sequence 9, Application US/08471969
; Patent No. 5871745
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOLY
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-969-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
```

```
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:|
Db             1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:||||||||||||| || ||| |||||||||: |||:|||||||||||||||||
Db            61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db           121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db           181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db           241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||    :|   |||||||||||||   ||   | | |||
Db           301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:||||||||| ||:|||||||:|||||||| ||||:||:|||:: |||||
Db           357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: ||||||:|||||||||||||||:|||:|||||:|||||||||||:
Db           416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| ||||||||:|||||||||||||||||||||:|:|||| |||||:||
Db           475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| ||||||||||||:: |||||||||||||||||||||||||||||:||||||
Db           535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:||||::|||||||||||||||||||||||||||||||
Db           595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|
Db           655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db           715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 30
US-08-384-137-9
; Sequence 9, Application US/08384137
; Patent No. 5871996
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-384-137-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:|
Db             1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:||||||||||||| || ||| |||||||||: |||:|||||||||||||||||
Db            61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db           121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db           181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db           241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||    :|   |||||||||||||   ||   | | |||
Db           301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:||||||||| ||:|||||||:|||||||| ||||:||:|||:: |||||
Db           357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: ||||||:|||||||||||||||:|||:|||||:|||||||||||:
Db           416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| ||||||||:|||||||||||||||||||||:|:|||| |||||:||
Db           475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| ||||||||||||:: |||||||||||||||||||||||||||||:||||||
Db           535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:||||::|||||||||||||||||||||||||||||||
```

```
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              |||||||||||: |||||||||| |:||||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 31
US-08-470-006A-9
; Sequence 9, Application US/08470006A
; Patent No. 5962217
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND
;    TITLE OF INVENTION:  BIOPOLYLMER CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,006A
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-470-006A-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
              |||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||||:::|||||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||| ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: ||||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:||||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              |||||||||||: |||||||||| |:||||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 32
US-08-691-563C-9
; Sequence 9, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
```

```
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|      | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|  |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             ||||||||:||: |||||:|||||||||||||:|||:|||||:|||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| |||||||||||:: |||||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: |||||||||| |:|||||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 33
US-09-200-990-9
```

```
; Sequence 9, Application US/09200990
; Patent No. 6184025
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/200,990
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-200-990-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|      | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|  |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356
```

```
Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 34
US-09-133-411-9
; Sequence 9, Application US/09133411
; Patent No. 6342383
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/133,411
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-133-411-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
Db          61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
Db         301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
Db         357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
Db         416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
Db         475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
Db         535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
Db         595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
Db         655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
Db         715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 35
US-09-374-766-9
; Sequence 9, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
```

```
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-374-766-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||    :|   ||||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:||||||||||||||||:|||:|||||:||||||||||||:

Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: ||||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 36
US-08-979-847B-9
; Sequence 9, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 9:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 741 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-979-847B-9
```

```
  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| |||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:||||||||||:|      | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:||||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| ||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 37
US-08-979-847B-112
; Sequence 112, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 112:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 758 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-08-979-847B-112

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
              ||||| |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
              ||||||||||:||    |||||| |||||:  |||||:|:||||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
              ||||| | ||||||||||:||||||:|||||||||||||| |: :||:||||||||||
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
              |||||||||||||| | |||||  |||||||||||||| |||||||||||||||| |:|||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
              ||||| |||||||  :|| | |  || |||||||||||||||| ||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
              ||||||    |||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| |||| ||||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
              |||||||||| ||||| |||||||| |||||  | || ||| ||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
```

```
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144

Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203

Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263

Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312

Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 38
US-08-471-724-8
; Sequence 8, Application US/08471724
; Patent No. 5800980
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,724
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-724-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967

Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027

Db          61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087

Db         121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147

Db         181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207

Db         241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267

Db         301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327

Db         361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387

Db         420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447

Db         477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507

Db         535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559

Db         594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 39
US-08-471-969-8
; Sequence 8, Application US/08471969
; Patent No. 5871745
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
```

```
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-969-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:|||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||:|:||    :::|||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 40
US-08-384-137-8
; Sequence 8, Application US/08384137
; Patent No. 5871996
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOLY
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-384-137-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:|||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||:|:||    :::|||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534
```

```
Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |:: ||||| |||||||||:||:|||||||:|||||||||| |||:|| |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 41
US-08-470-006A-8
; Sequence 8, Application US/08470006A
; Patent No. 5962217
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND
;    TITLE OF INVENTION:  BIOPOLYLMER CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,006A
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-470-006A-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
              ||||||||||:|||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
              ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
              |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
              |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
              ||||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
              :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
              || |||||||:  :|  ||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
              |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |:: ||||| |||||||||:||:|||||||:|||||||||| |||:|| |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 42
US-08-691-563C-8
; Sequence 8, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 base pairs
;      TYPE:  nucleotide
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:|||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 43
US-09-200-990-8
; Sequence 8, Application US/09200990
; Patent No. 6184025
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/200,990
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-200-990-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:|||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593
```

```
Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 44
US-09-133-411-8
; Sequence 8, Application US/09133411
; Patent No. 6342383
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/133,411
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-133-411-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  ||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || ||||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:||||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 45
US-09-374-766-8
; Sequence 8, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 base pairs
```

```
;       TYPE:  nucleotide
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-09-374-766-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :||||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::||||||||||:|||:|||:|||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | |  |||  |||||||||:|||||||||  ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 46
US-08-979-847B-8
; Sequence 8, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/979,847B
;            FILING DATE: 26-No. 6582703-1997
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 645 base pairs
;            TYPE: nucleotide
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-979-847B-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :||||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::||||||||||:|||:|||:|||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | |  |||  |||||||||:|||||||||  ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534
```

```
Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |:: ||||| |||||||||:||:|||||||:|||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||||||::||| ::||||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 47
US-08-979-847B-102
; Sequence 102, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 102:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 635 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-08-979-847B-102

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
              |||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:||||||| | || |||||||||||||  |||||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              ||||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              |||||||||||||||||||||||  |: |||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              ||||||||| |||||| ||||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | ||||||||||||||||||||||:||||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              ||||||||| |||||||||| |||||||||||||||||| |  |||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 48
US-08-686-878A-48
; Sequence 48, Application US/08686878A
; Patent No. 5708157
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Evans, Cheryl
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/686,878A
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
```

```
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  542 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-686-878A-48

  Query Match             6.8%;  Score 510.3;  DB 1;  Length 542;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  516;  Conservative   12;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
             ||||| || |||||||||||||||||| ||||| |||| |||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
             ||||::|||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
             | ||||| ||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
             |||||||||||||||||||||||||||||||||||| ||||||||:|||||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
             ||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
             ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
             |||
Db        540 CAT 542


RESULT 49
US-08-721-489-2
; Sequence 2, Application US/08721489
; Patent No. 5786465
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/721,489
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  542 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-721-489-2

  Query Match             6.8%;  Score 510.3;  DB 1;  Length 542;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  516;  Conservative   12;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
             ||||| || |||||||||||||||||| ||||| |||| |||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
             ||||::|||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
             | ||||| ||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
             |||||||||||||||||||||||||||||||||||| ||||||||:|||||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
             ||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
             ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
             |||
Db        540 CAT 542
```

RESULT 50
US-09-949-016-44302/c
; Sequence 44302, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44302
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44302

Query Match 6.2%; Score 467.1; DB 1; Length 601;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 507; Conservative 30; Mismatches 61; Indels 5; Gaps 2;

```
Qy       4059 TAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCC 4118
              ||  | |||: |||||||:||||||| |||||||||| :|:||| ||:|||||||| ||
Db        601 TAAAAAAAGGGAAAAGGGTAAATATACATACAGACTCAAAGTATTCTTACCTAGTCCTCT 542

Qy       4119 ATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCA 4178
              | |||||  ::|||||||:|:| |||||||||||||||||:|||:||||||||||||
Db        541 ACGCCCACACAGCAATATGGAAAAAAAGGGAATTCCTAACTTCTGAGGGAACACCTATCA 482

Qy       4179 MACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAG 4238
              :|||||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:||
Db        481 AACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAG 422

Qy       4239 TCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAG 4298
              ||||||||||:|||||||||| |||| ||||:|||| |||||:|||: ||:|||||||
Db        421 TCTTACACTGCCGGGGTCATCAGAAAGAAAAGGAAAGAAAAATAGAAGGAAACTGCCAAG 362

Qy       4299 CAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAAC 4358
              | : |||| |||| |||||||| |||||||| || |||||||| ||||||||| ||
Db        361 CGGGTATTAAAGCTAAAAGAGCCACAAGGCAGAACTCTCCATTAAAAATGCTTATAGAAG 302

Qy       4359 TTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAA 4418
              : ||| |||||| |||||||||| || ||||| |||||| | ||||||||||||||||
Db        301 KACCCCTAGTATGGGGTAATCCCCTCTGGGAACCCAAGCTGC--TACTCAGCAGGAGAAA 244

Qy       4419 CAGAATGGGGAACCTCAC---GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAG 4475
              |||| ||||||||||||   || | ||||| |||||||| | | || ||||||| |||
Db        243 TAGAACGGGGAACCTCACAAGGACGTAGTTTCCTCCCCTCAGAATGGCTAGCCACCGAAA 184

Qy       4476 AAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAA 4535
              |||| |||||||||||||||||||||  |||| |||||||||||||||||||||| ||||
Db        183 AAGGAAAAATACTTTTGCCTGCAACTAAACAATAGAAATTACTTAAAACCCTTCACCAAA 124

Qy       4536 CCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAG 4595
              |||||||||||||||  |||||||||||||:||||||||||||||||||||||||||||
Db        123 CCTTTCACTTAGGCATTAATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAC 64

Qy       4596 GCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATC 4655
              || |||||||||||| |||||||:|||:|||||||||||||:|||| ||:|:|||||| |
Db         63 GCTTTTTCAAAACTAACAAGCAGATAGTCAGGGCCTGTGAAGTGTGTCAAAAAAATAACC 4

Qy       4656 CCC 4658
              |||
Db          3 CCC 1
```

RESULT 51
US-09-949-016-44303/c
; Sequence 44303, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44303
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44303

Query Match 6.1%; Score 457.3; DB 1; Length 601;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 500; Conservative 59; Mismatches 41; Indels 7; Gaps 5;

```
Qy       2950 ACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTT 3009
              ||||||| ||:|::||||||||| ||| |||||||||| :||| |||| || |||||
Db        601 ACTTGAGACAGTTCTCATACCTGAACAGTCTTGTCCTTTGGTACATGGATAATCTACTTC 542

Qy       3010 TRGCYGCCYRTT-CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:|:|:|::||  | |||||||||||||||||||||||||||:|| ||:|| ||||||
Db        541 TAGCTGCCTGTTGAAAAAACCTTGTGCCATCAAGCCACCCAAGCGCTTTTAAACTTCCTC 482

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              |:|| ||||||||||:||||||||||:|:|:|||||:||||||| ||||||||||||
Db        481 ACCACTTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGTTAAA 422

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              |||||||:||||:|||||||||||||||:|||||||| || ||:::|||||||||
Db        421 TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTAAGAAATGTATCCAGCCT 362

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||||||:|||||||||||||||| ||| :| :|||||||:|||:|
Db        361 ATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAATAAAAG-GTTCCTTGGCATAACA 303

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |::|||||||  ||:|||||||  |||||||: |||:||||||||||:|||||:|||||:|
Db        302 GSCTTCTGCCAAATATGGATT-CCCAGGTACGGCAAAATAGCCAGGCCATTATATACACT 244

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||| |||||||||||||||||||||||:||| |||||: :|| |||:::|||
Db        243 AATTAAGGAAGCTCAGAAAGCCAATACCCATTTAGTAAAATGGACTCCTAAAGCAAAAGC 184

Qy       3369 GGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACR 3428
              |||||||||||||  ||| |||||||| |||||||| | ||||||||||:|| |||||:
Db        183 AGCTTTCCAGGCCCTAAAAAAAGGCCTT-AACCCAAGACTCAGTGTTAAGCTTACCAACG 125

Qy       3429 GGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTAC 3488
              |||||||||||||:||:|||::|||||   |||||||||:||:|||||| :|||||||||||
Db        124 GGGCAAGACTTTTCTTCATATGTCAC---AAAAAACAGAAATAGCTCCAGGAGTCCTTAC 68

Qy       3489 ACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT 3548
              ||||:| |:|||||: ||||||||| |::||||::|||| |:|||| |||:| |||||||
Db         67 ACAGGTTCAAGGGACCAGCTTGCAAACCATGGCATACCTAAGTAAGAAAATTAATGTAGT 8

Qy       3549 GGCAAAG 3555
              |||||||
Db          7 GGCAAAG 1
```

```
RESULT 52
US-09-513-999C-21394
; Sequence 21394, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 21394
;   LENGTH: 443
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 102
;   OTHER INFORMATION: k=g or t
US-09-513-999C-21394

  Query Match             5.8%;  Score 436.8;  DB 1;  Length 443;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  435;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         12 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 62

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db         63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTKGCAACTGCACTCTTCTGGTC 122

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 182

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 242

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:||||||||||||||||||||||||||||||||||||||||| ||||
Db        243 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 302

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        303 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 362

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        363 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 422

Qy        432 TCCATTCCTTGGAATCCGTGA 452
              |||||||||||||||||||||
Db        423 TCCATTCCTTGGAATCCGTGA 443


RESULT 53
US-08-979-847B-201
; Sequence 201, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 201:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 201:
US-08-979-847B-201

  Query Match             5.1%;  Score 382.5;  DB 1;  Length 438;
  Best Local Similarity   79.9%;  Pred. No. 0;
  Matches  350;  Conservative   82;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||:|:||||||:::||||||:|||:||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              :|:||: ||::|||||||||||||:||:|||||||||||||:||||||:||||||||:
Db         61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||:|:|||:::|||||||:||||||:::|||||||:|
Db        181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCST 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:||||||||||||:::::|||||||||:||::|
Db        241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:||||||||:||:|||||||  |||:|:|:  |:|||::|||||||::||||||:||||::
Db        301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
```

```
                |::|||||||||||||:|||||||:|||||||:||:|||||||||:|:||||:::::||||
Db        360 ADYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
                |||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 54
US-08-979-847B-128
; Sequence 128, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 128:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-08-979-847B-128

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  375;  Conservative   45;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
                ||||||||:||:|::|||||||||||||:|||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
                ||:||: ||::||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
                ||:|||||||||||||:|||||||||:|:|:||||||:||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
                ||||||||:||||:||||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
                ||||||| ||||||||||:||||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
                ||:|||||| ||:|||||| |||||||:  ||:|::|||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
                |||||||||||||||| |||||||:||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
                |||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 55
US-08-979-847B-129
; Sequence 129, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 129:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 129:
US-08-979-847B-129
```

```
  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:| |||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 56
US-08-979-847B-131
; Sequence 131, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 131:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 131:
US-08-979-847B-131

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:|||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 57
US-08-979-847B-130
; Sequence 130, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
```

```
;              STREET: P.O. BOX 19928
;              CITY: ALEXANDRIA
;              STATE: VA
;              COUNTRY: USA
;              ZIP: 22320
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/08/979,847B
;              FILING DATE: 26-No. 6582703-1997
;              CLASSIFICATION: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: BERRIDGE, WILLIAM P.
;              REGISTRATION NUMBER: 30,024
;              REFERENCE/DOCKET NUMBER: WPB 39046A
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-836-6400
;              TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 130:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 438 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         SEQUENCE DESCRIPTION: SEQ ID NO: 130:
US-08-979-847B-130

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::||||||||||||||||||||||||||||||:||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  ||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||| ||||||| ||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 58
US-08-979-847B-138
; Sequence 138, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 138:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-08-979-847B-138

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| || ||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:||||||:||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |||||:|||||||| ||||||:::|||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||:::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||||| |:  ||:|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350
```

```
Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  |||||||||||||||||||||||||||||| ||:||||||||:|:||||| :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 59
US-08-979-847B-135
; Sequence 135, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 135:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-08-979-847B-135

  Query Match             4.6%;  Score 344.6;  DB 1;  Length 429;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  347;  Conservative   60;  Mismatches   21;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| |:||:|||:: ||||||:||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCYTCAGTATRGGGATGAYTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| ||:|||||||||||::|||||:||||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          |||||||  |:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAARGGCTCASCTCTGCTCACASCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |:|||:|||||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:|||||||||||||::::||||||||||:||::|
Db        232 ATACTGGYTTATCCYCATCCCAWAACCYTAAAGCAACTAAGARGGTTCCTTGGCATAWCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||||||: | |:  |::|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCV-GATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||||||||||||||||||| ||:|||||||||:|:||||: :::||||
Db        351 ADYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGARACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 60
US-08-979-847B-136
; Sequence 136, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 136:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
```

```
;       SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-08-979-847B-136

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| ||||||||||||||||:|||||:||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:|:||||| |||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:|||||||| |||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:|| ||||||||||||||||:|::|||||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||| | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 61
US-09-880-107-538/c
; Sequence 538, Application US/09880107
; Patent No. 6974667
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 538
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. 6974667 AA250958
US-09-880-107-538

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
              ||||| |||||||||||||||||||| |||: || ||||||||| ||||||||||||

Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
              ||||||||||||||  |||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
              ||||||||:||| |||||:|||||||||||||| ||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
              |||||:|||||||  ||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
              |||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| |  |||||||| || |||||||||| |||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||||  |||||||||||||||||||||||:|||
Db          50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 62
US-08-979-847B-137
; Sequence 137, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 137:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
```

```
;           TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-08-979-847B-137

  Query Match             4.5%;  Score 339;  DB 1;  Length 429;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  358;  Conservative   40;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:  |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATAGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| |||||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |: :|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAGGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| | |||:|||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:|||||||| |||||||||||||:|::|||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||||  | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              | |||||||||||||||||||||||||||||| ||:|||||||||:|:|||| :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 63
US-08-686-878A-47
; Sequence 47, Application US/08686878A
; Patent No. 5708157
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Evans, Cheryl
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/686,878A
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  374 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-686-878A-47

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:|||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              |  |||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              |  |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 64
US-08-721-489-1
; Sequence 1, Application US/08721489
; Patent No. 5786465
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/721,489
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  374 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-721-489-1

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:|||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             |||||||||||||||||||||||| |||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             |  |||||||||||||||||:| |||||||||||| ||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             |  |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 65
US-08-686-878A-50
; Sequence 50, Application US/08686878A
; Patent No. 5708157
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Evans, Cheryl
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/686,878A
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  279 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-686-878A-50

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| |||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
              :|||||||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db         121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db         241 TNTATTAAATNTTGCAACTGCA 262


RESULT 66
US-08-721-489-4
; Sequence 4, Application US/08721489
; Patent No. 5786465
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/721,489
;       FILING DATE:
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brown, Scott A.
;       REGISTRATION NUMBER:  32,724
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 498-8224
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  279 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-721-489-4

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
             :||||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 67
US-08-979-847B-100
; Sequence 100, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/979,847B
;            FILING DATE: 26-No. 6582703-1997
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 100:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 310 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 100:
US-08-979-847B-100

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             ||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||     | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||||  |||||| ||| ||||  |||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             |||||| |||||||| |||||||||||| ||||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             |||||||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 68
US-08-471-724-2
; Sequence 2, Application US/08471724
; Patent No. 5800980
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOLY
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,724
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-724-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| ||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 69
US-08-471-969-2
; Sequence 2, Application US/08471969
; Patent No. 5871745
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-969-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| ||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 70
US-08-384-137-2
; Sequence 2, Application US/08384137
; Patent No. 5871996
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,137
```

```
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-384-137-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  ||||||||  ||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| ||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 71
US-08-470-006A-2
; Sequence 2, Application US/08470006A
; Patent No. 5962217
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND
;    TITLE OF INVENTION:  BIOPOLYLMER CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,006A
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-470-006A-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  ||||||||  ||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| ||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 72
US-08-691-563C-2
; Sequence 2, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
```

```
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||  ||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             ||||  |||||  |||||  | |  ||||||||||||||||||  |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:||||||||||  |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||||  |||||||||||||||||||||||||||||||||  |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||||:||||||||||||||||  ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 73
US-09-200-990-2
; Sequence 2, Application US/09200990
; Patent No. 6184025
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/200,990
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
```

```
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-200-990-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||  ||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             ||||  |||||  |||||  | |  ||||||||||||||||||  |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:||||||||||  |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||||  |||||||||||||||||||||||||||||||||  |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||||:||||||||||||||||  ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 74
US-09-133-411-2
; Sequence 2, Application US/09133411
; Patent No. 6342383
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/133,411
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
```

```
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-133-411-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 75
US-09-374-766-2
; Sequence 2, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-374-766-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 76
US-08-979-847B-2
; Sequence 2, Application US/08979847B
; Patent No. 6582703
;  GENERAL INFORMATION:
;      APPLICANT: PERRON, HERVE
;                 BESEME, FREDERIC
;                 BEDIN, FREDERIC
;                 PARANHOS-BACCALA, GLAUCIA
;                 KOMURIAN-PRADEL, FLORENCE
;                 JOLIVET-REYNAUD, COLETTE
;                 MANDRAND, BERNARD
;                 GARSON, JEREMY
;                 TUKE, PHILIP
;      TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                          ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                          THERAPEUTIC PURPOSES
;      NUMBER OF SEQUENCES: 210
;      CORRESPONDENCE ADDRESS:
;           ADDRESSEE: OLIFF & BERRIDGE, PLC
;           STREET: P.O. BOX 19928
;           CITY: ALEXANDRIA
;           STATE: VA
;           COUNTRY: USA
;           ZIP: 22320
;      COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/08/979,847B
;           FILING DATE: 26-No. 6582703-1997
```

```
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 297 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-979-847B-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 77
US-08-979-847B-36
; Sequence 36, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 36:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 150 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-08-979-847B-36

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  ||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              ||||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 78
US-08-471-724-7
; Sequence 7, Application US/08471724
; Patent No. 5800980
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,724
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
```

```
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-724-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 79
US-08-471-969-7
; Sequence 7, Application US/08471969
; Patent No. 5871745
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-471-969-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 80
US-08-384-137-7
; Sequence 7, Application US/08384137
; Patent No. 5871996
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-384-137-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 81
US-08-470-006A-7
; Sequence 7, Application US/08470006A
; Patent No. 5962217
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
```

```
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND
;    TITLE OF INVENTION:  BIOPOLYLMER CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,006A
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-470-006A-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
                :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
                :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 82
US-08-691-563C-7
; Sequence 7, Application US/08691563C
; Patent No. 6001987
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563C
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-691-563C-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
                :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
                :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 83
US-09-200-990-7
; Sequence 7, Application US/09200990
; Patent No. 6184025
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/200,990
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/471,969
;      FILING DATE:  June 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
```

```
;      REFERENCE/DOCKET NUMBER:  WPB 36055A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-200-990-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 84
US-09-133-411-7
; Sequence 7, Application US/09133411
; Patent No. 6342383
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Francois MALLET
;    APPLICANT:  Bernard MANDRAND
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Frederic BESEME
;    TITLE OF INVENTION:  MSRV1 VIRUS AND MSRV2 PATHOGEN AND/OR
;    TITLE OF INVENTION:  INFECTIVE AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, AND BIOPOL
;    TITLE OF INVENTION:  CONSTITUENTS THEREOF (AS AMENDED)
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/133,411
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/384,137
;      FILING DATE:  February 6, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 36055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-133-411-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 85
US-09-374-766-7
; Sequence 7, Application US/09374766
; Patent No. 6579526
;  GENERAL INFORMATION:
;    APPLICANT:  Herve PERRON
;    APPLICANT:  Frederic BESEME
;    APPLICANT:  Frederic BEDIN
;    APPLICANT:  Glaucia PARANHOS-BACCALA
;    APPLICANT:  Florence KOMURIAN-PRADEL
;    APPLICANT:  Colette JOLIVET
;    APPLICANT:  Bernard MANDRAND
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oliff & Berridge
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,766
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,563
;      FILING DATE:  02-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 38588
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-374-766-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
```

```
              :|::||||:|::|:|:::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 86
US-08-979-847B-7
; Sequence 7, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 111 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-847B-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


Search completed: August  7, 2006, 16:28:47
Job time : 494 secs
```

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2006, 16:43:42 ; Search time 1840 Seconds
(without alignments)
2.657 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 7489.6
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 425 seqs, 322401 residues

Total number of hits satisfying chosen parameters: 849

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

N. Genesoq

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 424 summaries

Database : ngsdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 7487.42 | 99.9 | 7582 | 1 | AAX25665 | Complete human end |
| 2 | 3549.5 | 47.4 | 4349 | 1 | AAS76475 | DNA encoding novel |
| 3 | 3025.7 | 40.4 | 3464 | 1 | ADE09587 | Novel DNA-related |
| 4 | 2890.1 | 38.6 | 2930 | 1 | AAD24195 | Human syncytin cDN |
| 5 | 2889.1 | 38.6 | 2946 | 1 | AAX77526 | Human secreted pro |
| 6 | 2889.1 | 38.6 | 2946 | 1 | AAZ59468 | Human secreted pro |
| 7 | 2889.1 | 38.6 | 2946 | 1 | ADC38776 | Human cDNA encodin |
| 8 | 2750.9 | 36.7 | 3831 | 1 | AAS71727 | DNA encoding novel |
| 9 | 2706.61 | 36.1 | 2938 | 1 | AAX25656 | Human endogenous r |
| 10 | 2706.61 | 36.1 | 2938 | 1 | AAA59206 | Gag and partial po |
| 11 | 2626.2 | 35.1 | 3372 | 1 | AAX25663 | Human endogenous r |
| 12 | 2626.2 | 35.1 | 3372 | 1 | AAA59213 | Partial pol gene a |
| 13 | 2551.4 | 34.1 | 2599 | 1 | ABN97927 | Human retroviral s |
| 14 | 2315.4 | 30.9 | 2784 | 1 | ABN97930 | Human retroviral s |
| 15 | 2256.8 | 30.1 | 2372 | 1 | AAX25664 | Human endogenous r |
| 16 | 2256.8 | 30.1 | 2372 | 1 | AAA59214 | R-U5 region and pa |
| 17 | 2205.1 | 29.4 | 2782 | 1 | AAX25661 | Human endogenous r |
| 18 | 2205.1 | 29.4 | 2782 | 1 | AAA59211 | 5' non coding, 3' |
| 19 | 2205.1 | 29.4 | 2782 | 1 | AAH20069 | HERV-W envelope pr |
| 20 | 2201.9 | 29.4 | 2781 | 1 | AAF55630 | Nucleotide sequenc |
| 21 | 2141.4 | 28.6 | 2575 | 1 | AAX25668 | Human endogenous r |
| 22 | 2036.1 | 27.2 | 2585 | 1 | AAX25667 | Human endogenous r |
| 23 | 2032.5 | 27.1 | 2055 | 1 | ABN97948 | Human retroviral H |
| 24 | 1851.9 | 24.7 | 2304 | 1 | AAV43199 | Multiple sclerosis |
| 25 | 1851.9 | 24.7 | 2304 | 1 | ADG14828 | MSRV associated cD |
| 26 | 1848.6 | 24.7 | 2006 | 1 | AAX25658 | Human endogenous r |
| 27 | 1848.6 | 24.7 | 2006 | 1 | AAA59208 | 3' pol gene and 5' |
| 28 | 1847.5 | 24.7 | 2304 | 1 | ADB84382 | MSRV-1 gene encodi |
| 29 | 1812.4 | 24.2 | 2074 | 1 | AAD41225 | Human EMBRY-2 cDNA |
| 30 | 1780.4 | 23.8 | 2030 | 1 | AAA63826 | Nucleotide sequenc |
| 31 | 1607 | 21.5 | 1617 | 1 | AAH20070 | HERV-W envelope pr |
| 32 | 1607 | 21.5 | 1617 | 1 | AED10942 | HERV envelope prot |
| 33 | 1601 | 21.4 | 1614 | 1 | AED10984 | HERV R393E/Q envel |
| 34 | 1598 | 21.3 | 1614 | 1 | AED10986 | HERV R393E/Q F399A |
| 35 | 1551.5 | 20.7 | 1894 | 1 | AAI14608 | Probe #4541 for ge |
| 36 | 1551.5 | 20.7 | 1894 | 1 | ABA56337 | Human foetal liver |
| 37 | 1551.5 | 20.7 | 1894 | 1 | AAI35980 | Probe #4666 used t |
| 38 | 1551.5 | 20.7 | 1894 | 1 | ABA45822 | Human breast cell |
| 39 | 1551.5 | 20.7 | 1894 | 1 | ABA25978 | Probe #4444 for ge |
| 40 | 1551.5 | 20.7 | 1894 | 1 | AAK30018 | Human bone marrow |
| 41 | 1551.5 | 20.7 | 1894 | 1 | AAK04516 | Human brain expres |
| 42 | 1551.5 | 20.7 | 1894 | 1 | ABS29670 | Human liver single |
| 43 | 1551.5 | 20.7 | 1894 | 1 | AAI04422 | Probe #4413 used t |
| 44 | 1551.5 | 20.7 | 1894 | 1 | ABS04589 | Human genome-deriv |
| c 45 | 1487.1 | 19.9 | 1799 | 1 | ABN97931 | Human retroviral s |
| 46 | 1440.3 | 19.2 | 1948 | 1 | AAX25659 | Human endogenous r |
| 47 | 1440.3 | 19.2 | 1948 | 1 | AAA59209 | 5' non coding, 3' |
| 48 | 1431 | 19.1 | 1629 | 1 | AAA96625 | DNA encoding an en |
| 49 | 1383.7 | 18.5 | 1715 | 1 | AAS73701 | DNA encoding novel |
| 50 | 1329 | 17.7 | 1422 | 1 | AAX25657 | Human endogenous r |
| 51 | 1329 | 17.7 | 1422 | 1 | AAA59207 | Fragment 3' of the |
| 52 | 1324.2 | 17.7 | 1481 | 1 | AAV43217 | Multiple sclerosis |
| 53 | 1324.2 | 17.7 | 1481 | 1 | AAX29703 | Clone C15 from MSR |
| 54 | 1324.2 | 17.7 | 1481 | 1 | ADG14846 | MSRV associated cD |
| 55 | 1321.4 | 17.6 | 1600 | 1 | ADB84502 | MSRV-1 pol gene co |
| 56 | 1320.4 | 17.6 | 1600 | 1 | ADG14948 | MSRV associated cD |
| 57 | 1309.8 | 17.5 | 1600 | 1 | ADG14947 | MSRV associated cD |
| 58 | 1309.2 | 17.5 | 1600 | 1 | ADB84501 | MSRV-1 pol gene co |
| 59 | 1296.4 | 17.3 | 1481 | 1 | ADB84400 | MSRV-1 associated |
| 60 | 1291.9 | 17.2 | 1597 | 1 | ADG14946 | MSRV associated cD |
| 61 | 1288.1 | 17.2 | 1597 | 1 | ADB84500 | MSRV-1 pol gene. |
| 62 | 1285.7 | 17.2 | 1660 | 1 | AAS76459 | DNA encoding novel |
| 63 | 1285.7 | 17.2 | 1660 | 1 | AAS73608 | DNA encoding novel |
| 64 | 1285.7 | 17.2 | 1660 | 1 | AAS84201 | DNA encoding novel |
| 65 | 1270.4 | 17.0 | 1321 | 1 | AAX25655 | Human endogenous r |
| 66 | 1270.4 | 17.0 | 1321 | 1 | AAA59205 | 5' non coding and |
| 67 | 1229.2 | 16.4 | 1472 | 1 | AAI19284 | Probe #9217 for ge |
| 68 | 1229.2 | 16.4 | 1472 | 1 | ABA64295 | Human foetal liver |
| 69 | 1229.2 | 16.4 | 1472 | 1 | AAI44454 | Probe #13140 used |
| 70 | 1229.2 | 16.4 | 1472 | 1 | ABS38070 | Human liver single |
| 71 | 1229.2 | 16.4 | 1472 | 1 | ABS12550 | Human genome-deriv |
| 72 | 1201.4 | 16.0 | 1326 | 1 | ABN97928 | Human retroviral s |
| 73 | 1117.7 | 14.9 | 1511 | 1 | AAV43228 | Multiple sclerosis |
| 74 | 1117.7 | 14.9 | 1511 | 1 | AAX29706 | Clone CL2 from MSR |
| 75 | 1117.7 | 14.9 | 1511 | 1 | AAA59188 | Endogenetic retrov |
| 76 | 1117.7 | 14.9 | 1511 | 1 | ADG14858 | MSRV associated cD |
| 77 | 1106.8 | 14.8 | 1136 | 1 | AAX25660 | Human endogenous r |
| 78 | 1106.8 | 14.8 | 1136 | 1 | AAA59210 | 3' pol gene and 3' |
| 79 | 1093.5 | 14.6 | 1511 | 1 | ADB84412 | MSRV-1 associated |
| 80 | 960.2 | 12.8 | 1080 | 1 | ABN97949 | Human retroviral H |
| 81 | 952.8 | 12.7 | 1167 | 1 | AAT96476 | MSRV-1 clone FP6. |
| 82 | 952.8 | 12.7 | 1167 | 1 | AAV43170 | Multiple sclerosis |
| 83 | 952.8 | 12.7 | 1167 | 1 | ADB84352 | MSRV-1 pol/RT regi |
| 84 | 952.8 | 12.7 | 1167 | 1 | ADG14798 | MSRV-1 pol ORF #1. |
| 85 | 948 | 12.7 | 1117 | 1 | AAI28063 | Probe #17996 for g |
| 86 | 948 | 12.7 | 1117 | 1 | ABA76381 | Human foetal liver |
| 87 | 948 | 12.7 | 1117 | 1 | AAI57054 | Probe #25740 used |
| 88 | 948 | 12.7 | 1117 | 1 | ABS50596 | Human liver single |
| 89 | 948 | 12.7 | 1117 | 1 | ABS24539 | Human genome-deriv |
| 90 | 948 | 12.7 | 1117 | 1 | ACH90651 | Human genome deriv |
| 91 | 936.9 | 12.5 | 1158 | 1 | AAT02815 | Multiple sclerosis |
| 92 | 936.9 | 12.5 | 1158 | 1 | AAT37973 | Partial sequence o |
| 93 | 936.9 | 12.5 | 1158 | 1 | AAT96527 | MSRV-1B pol* seque |
| 94 | 936.9 | 12.5 | 1158 | 1 | AAV43115 | Multiple sclerosis |
| 95 | 936.9 | 12.5 | 1158 | 1 | ADB84549 | MSRV-1 pol gene co |
| 96 | 936.9 | 12.5 | 1158 | 1 | ADG14742 | MSRV-1 pol gene. |
| c 97 | 894.401 | 11.9 | 1216 | 1 | ABN97933 | Human retroviral s |
| 98 | 714.8 | 9.5 | 910 | 1 | AAS71725 | DNA encoding novel |
| 99 | 712.9 | 9.5 | 792 | 1 | AAI23803 | Probe #13736 for g |
| 100 | 712.9 | 9.5 | 792 | 1 | ABA68919 | Human foetal liver |
| 101 | 712.9 | 9.5 | 792 | 1 | AAI49115 | Probe #17801 used |
| 102 | 712.9 | 9.5 | 792 | 1 | ABA50940 | Human breast cell |
| 103 | 712.9 | 9.5 | 792 | 1 | ABA35872 | Probe #14338 for g |
| 104 | 712.9 | 9.5 | 792 | 1 | AAK43040 | Human bone marrow |
| 105 | 712.9 | 9.5 | 792 | 1 | AAK17251 | Human brain expres |
| 106 | 712.9 | 9.5 | 792 | 1 | ABS42675 | Human liver single |

| | | | | | | |
|---|---|---|---|---|---|---|
| 107 | 712.9 | 9.5 | 792 | 1 | AAI09412 | Probe #9403 used t |
| 108 | 712.9 | 9.5 | 792 | 1 | ABS17105 | Human genome-deriv |
| 109 | 674.5 | 9.0 | 944 | 1 | AAI18958 | Probe #8891 for ge |
| 110 | 674.5 | 9.0 | 944 | 1 | AAI44090 | Probe #12776 used |
| 111 | 674.5 | 9.0 | 944 | 1 | ABS37794 | Human liver single |
| 112 | 674.5 | 9.0 | 944 | 1 | ABS12201 | Human genome-deriv |
| 113 | 662.8 | 8.8 | 924 | 1 | AAS72224 | DNA encoding novel |
| c 114 | 646 | 8.6 | 902 | 1 | AAS71722 | DNA encoding novel |
| 115 | 637.7 | 8.5 | 841 | 1 | AAS76467 | DNA encoding novel |
| 116 | 586 | 7.8 | 741 | 1 | AAT02823 | Multiple sclerosis |
| 117 | 586 | 7.8 | 741 | 1 | AAT37981 | Rheumatoid polyart |
| 118 | 586 | 7.8 | 741 | 1 | AAT96517 | Clone PSJ17 compri |
| 119 | 586 | 7.8 | 741 | 1 | AAV43123 | Multiple sclerosis |
| 120 | 586 | 7.8 | 741 | 1 | ADB84557 | MSRV-1DNA clone PS |
| 121 | 586 | 7.8 | 741 | 1 | ADG14750 | MSRV-1 clone PSJ17 |
| 122 | 564.8 | 7.5 | 758 | 1 | AAV43222 | Multiple sclerosis |
| 123 | 564.8 | 7.5 | 758 | 1 | AAX29705 | Clone LB16 from MS |
| 124 | 564.8 | 7.5 | 758 | 1 | ADG14853 | MSRV associated cD |
| 125 | 560.3 | 7.5 | 745 | 1 | AAH04815 | Human cDNA clone ( |
| 126 | 555.4 | 7.4 | 758 | 1 | ADB84407 | MSRV-1 associated |
| 127 | 533.4 | 7.1 | 648 | 1 | AAT37983 | Rheumatoid arthrit |
| 128 | 533.4 | 7.1 | 648 | 1 | AAT37984 | Rheumatoid arthrit |
| 129 | 523.9 | 7.0 | 645 | 1 | AAT37980 | MSRV-1B POL gene f |
| 130 | 523.9 | 7.0 | 645 | 1 | AAT96523 | MSRV-1B clone M003 |
| 131 | 523.9 | 7.0 | 645 | 1 | AAV43122 | Multiple sclerosis |
| 132 | 523.9 | 7.0 | 645 | 1 | ADB84556 | MSRV-1 DNA clone M |
| 133 | 523.9 | 7.0 | 645 | 1 | ADG14749 | MSRV-1 clone PSJ17 |
| 134 | 521.1 | 7.0 | 645 | 1 | AAT02822 | Multiple sclerosis |
| c 135 | 520.3 | 6.9 | 591 | 1 | ACH89452 | Human genome deriv |
| 136 | 517.4 | 6.9 | 635 | 1 | AAV43215 | Multiple sclerosis |
| 137 | 517.4 | 6.9 | 635 | 1 | AAX29702 | Clone CL6-3' from |
| 138 | 517.4 | 6.9 | 635 | 1 | ADG14843 | MSRV associated cD |
| 139 | 516.8 | 6.9 | 591 | 1 | AAX25685 | Human endogenous r |
| 140 | 516.8 | 6.9 | 591 | 1 | AAA59204 | Probe Penv-C15 use |
| 141 | 510.3 | 6.8 | 542 | 1 | AAV09124 | Nucleotide sequenc |
| 142 | 510.3 | 6.8 | 542 | 1 | AAV15134 | Human adult testes |
| 143 | 510 | 6.8 | 635 | 1 | ADB84397 | MSRV-1 associated |
| 144 | 505.2 | 6.7 | 678 | 1 | AAX25683 | Human endogenous r |
| 145 | 505.2 | 6.7 | 678 | 1 | AAA59203 | Probe Ppol-MSRV us |
| 146 | 489.8 | 6.5 | 583 | 1 | AAS88391 | DNA encoding novel |
| c 147 | 481.1 | 6.4 | 549 | 1 | ABA61335 | Human foetal liver |
| c 148 | 481.1 | 6.4 | 549 | 1 | AAI41239 | Probe #9925 used t |
| c 149 | 481.1 | 6.4 | 549 | 1 | ABA29141 | Probe #7607 for ge |
| c 150 | 481.1 | 6.4 | 549 | 1 | AAK35525 | Human bone marrow |
| c 151 | 481.1 | 6.4 | 549 | 1 | AAK09633 | Human brain expres |
| c 152 | 481.1 | 6.4 | 549 | 1 | ABS35255 | Human liver single |
| c 153 | 481.1 | 6.4 | 549 | 1 | ABS09860 | Human genome-deriv |
| c 154 | 477.3 | 6.4 | 559 | 1 | ACH75729 | Human genome deriv |
| c 155 | 473.9 | 6.3 | 568 | 1 | ACH67368 | Human genome deriv |
| 156 | 460.7 | 6.2 | 568 | 1 | AAI44528 | Probe #13214 used |
| c 157 | 454 | 6.1 | 548 | 1 | ABA62059 | Human foetal liver |
| c 158 | 454 | 6.1 | 548 | 1 | AAI41998 | Probe #10684 used |
| c 159 | 454 | 6.1 | 548 | 1 | AAK36276 | Human bone marrow |
| c 160 | 454 | 6.1 | 548 | 1 | AAK10373 | Human brain expres |
| c 161 | 454 | 6.1 | 548 | 1 | ABS35955 | Human liver single |
| c 162 | 454 | 6.1 | 548 | 1 | ABS10362 | Human genome-deriv |
| 163 | 448.6 | 6.0 | 578 | 1 | ACH77079 | Human genome deriv |
| c 164 | 445.1 | 5.9 | 559 | 1 | AAI17000 | Probe #6933 for ge |
| c 165 | 445.1 | 5.9 | 559 | 1 | ABA61156 | Human foetal liver |
| c 166 | 445.1 | 5.9 | 559 | 1 | AAI41056 | Probe #9742 used t |
| c 167 | 445.1 | 5.9 | 559 | 1 | ABA29035 | Probe #7501 for ge |
| c 168 | 445.1 | 5.9 | 559 | 1 | AAK35343 | Human bone marrow |
| c 169 | 445.1 | 5.9 | 559 | 1 | AAK09451 | Human brain expres |
| c 170 | 445.1 | 5.9 | 559 | 1 | ABS35081 | Human liver single |
| c 171 | 445.1 | 5.9 | 559 | 1 | ABS09734 | Human genome-deriv |
| 172 | 436.8 | 5.8 | 443 | 1 | AAC17319 | Human secreted pro |
| 173 | 433.2 | 5.8 | 491 | 1 | AAI49732 | Probe #18418 used |
| 174 | 429.5 | 5.7 | 566 | 1 | ACH87164 | Human genome deriv |
| 175 | 423 | 5.6 | 507 | 1 | ACH73335 | Human genome deriv |
| 176 | 420.9 | 5.6 | 558 | 1 | AAI44191 | Probe #12877 used |
| c 177 | 420.2 | 5.6 | 521 | 1 | ACH71117 | Human genome deriv |
| 178 | 409.6 | 5.5 | 488 | 1 | AAI11856 | Probe #1789 for ge |
| 179 | 409.6 | 5.5 | 488 | 1 | ABA53554 | Human foetal liver |
| 180 | 409.6 | 5.5 | 488 | 1 | AAI33179 | Probe #1865 used t |
| 181 | 409.6 | 5.5 | 488 | 1 | ABA43137 | Human breast cell |
| 182 | 409.6 | 5.5 | 488 | 1 | ABA23311 | Probe #1777 for ge |
| 183 | 409.6 | 5.5 | 488 | 1 | AAK27280 | Human bone marrow |
| 184 | 409.6 | 5.5 | 488 | 1 | AAK01823 | Human brain expres |
| 185 | 409.6 | 5.5 | 488 | 1 | ABS26854 | Human liver single |
| 186 | 409.6 | 5.5 | 488 | 1 | AAI01794 | Probe #1785 used t |
| 187 | 409.6 | 5.5 | 488 | 1 | ABS01806 | Human genome-deriv |
| c 188 | 400.4 | 5.3 | 504 | 1 | ABA60281 | Human foetal liver |
| c 189 | 400.4 | 5.3 | 504 | 1 | AAI40167 | Probe #8853 used t |
| c 190 | 400.4 | 5.3 | 504 | 1 | ABA28561 | Probe #7027 for ge |
| c 191 | 400.4 | 5.3 | 504 | 1 | AAK34442 | Human bone marrow |
| c 192 | 400.4 | 5.3 | 504 | 1 | AAK08559 | Human brain expres |
| c 193 | 400.4 | 5.3 | 504 | 1 | ABS34221 | Human liver single |
| c 194 | 400.4 | 5.3 | 504 | 1 | ABS09114 | Human genome-deriv |
| 195 | 387.5 | 5.2 | 474 | 1 | ACH90788 | Human genome deriv |
| 196 | 382.5 | 5.1 | 438 | 1 | ADG14942 | MSRV associated cD |
| 197 | 381.3 | 5.1 | 438 | 1 | AAV43239 | MSRV-1 clone 43 po |
| 198 | 381.3 | 5.1 | 438 | 1 | AAV43237 | MSRV-1 clone 41 po |
| 199 | 381.3 | 5.1 | 438 | 1 | ADB84424 | MSRV-1 pol region |
| 200 | 381.3 | 5.1 | 438 | 1 | ADB84426 | MSRV-1 pol region |
| 201 | 381.3 | 5.1 | 438 | 1 | ADB84423 | MSRV-1 pol region |
| 202 | 381.3 | 5.1 | 438 | 1 | ADG14872 | MSRV cDNA #4. Mul |
| 203 | 381.3 | 5.1 | 438 | 1 | ADG14870 | MSRV cDNA #2. Mul |
| 204 | 381.3 | 5.1 | 438 | 1 | ADG14869 | MSRV cDNA #1. Mul |
| 205 | 379.7 | 5.1 | 438 | 1 | AAV43238 | MSRV-1 clone 42 po |
| 206 | 379.7 | 5.1 | 438 | 1 | ADB84425 | MSRV-1 pol region |
| 207 | 379.7 | 5.1 | 438 | 1 | ADG14871 | MSRV cDNA #3. Mul |
| 208 | 378.6 | 5.1 | 433 | 1 | AAI36615 | Probe #5301 used t |
| 209 | 377.7 | 5.0 | 422 | 1 | AAI13992 | Probe #3925 for ge |
| 210 | 377.7 | 5.0 | 422 | 1 | AAI35373 | Probe #4059 used t |
| 211 | 377.7 | 5.0 | 422 | 1 | ABA45226 | Human breast cell |
| 212 | 377.7 | 5.0 | 422 | 1 | AAK29413 | Human bone marrow |
| 213 | 377.7 | 5.0 | 422 | 1 | ABS29038 | Human liver single |
| 214 | 376.9 | 5.0 | 438 | 1 | ADB84496 | MSRV-1 associated |
| c 215 | 357.2 | 4.8 | 448 | 1 | AAS87567 | DNA encoding novel |
| c 216 | 353.6 | 4.7 | 439 | 1 | AAI63993 | Human polynucleoti |
| c 217 | 353.6 | 4.7 | 439 | 1 | AAS31671 | Genomic sequence # |
| c 218 | 353.6 | 4.7 | 439 | 1 | ADM24708 | Human PRO polynucl |
| 219 | 352.9 | 4.7 | 384 | 1 | AAV86331 | EST clone AJ172. |
| 220 | 345 | 4.6 | 429 | 1 | AAV43243 | MSRV-1 clone 8 pol |
| 221 | 345 | 4.6 | 429 | 1 | ADB84430 | MSRV-1 pol region |
| 222 | 345 | 4.6 | 429 | 1 | ADG14879 | MSRV cDNA #8. Mul |
| 223 | 344.6 | 4.6 | 429 | 1 | ADB84433 | MSRV-1 pol region |
| 224 | 341 | 4.6 | 429 | 1 | AAV43241 | MSRV-1 clone 1 pol |
| 225 | 341 | 4.6 | 429 | 1 | ADB84431 | MSRV-1 pol region |
| 226 | 341 | 4.6 | 429 | 1 | ADG14877 | MSRV cDNA #6. Mul |
| 227 | 339.4 | 4.5 | 429 | 1 | ADG14878 | MSRV cDNA #7. Mul |
| 228 | 339.4 | 4.5 | 429 | 1 | ADG14876 | MSRV cDNA #5. Mul |
| c 229 | 339 | 4.5 | 410 | 1 | ABK64210 | Human benign prost |
| c 230 | 339 | 4.5 | 410 | 1 | ABN94040 | Gene #538 used to |
| 231 | 339 | 4.5 | 429 | 1 | AAV43242 | MSRV-1 clone 5 pol |
| 232 | 339 | 4.5 | 429 | 1 | ADB84432 | MSRV-1 pol region |
| c 233 | 335.8 | 4.5 | 405 | 1 | AAI13454 | Probe #3387 for ge |
| c 234 | 335.8 | 4.5 | 405 | 1 | ABA55151 | Human foetal liver |
| c 235 | 335.8 | 4.5 | 405 | 1 | AAI34810 | Probe #3496 used t |
| c 236 | 335.8 | 4.5 | 405 | 1 | ABA44696 | Human breast cell |
| c 237 | 335.8 | 4.5 | 405 | 1 | ABA24898 | Probe #3364 for ge |
| c 238 | 335.8 | 4.5 | 405 | 1 | AAK28864 | Human bone marrow |
| c 239 | 335.8 | 4.5 | 405 | 1 | AAK03413 | Human brain expres |
| c 240 | 335.8 | 4.5 | 405 | 1 | ABS28480 | Human liver single |
| c 241 | 335.8 | 4.5 | 405 | 1 | AAI03334 | Probe #3325 used t |
| c 242 | 335.8 | 4.5 | 405 | 1 | ABS03396 | Human genome-deriv |
| c 243 | 324.7 | 4.3 | 409 | 1 | AAI14299 | Probe #4232 for ge |
| c 244 | 324.7 | 4.3 | 409 | 1 | ABA56026 | Human foetal liver |
| c 245 | 324.7 | 4.3 | 409 | 1 | AAI35675 | Probe #4361 used t |
| c 246 | 324.7 | 4.3 | 409 | 1 | ABA45530 | Human breast cell |
| c 247 | 324.7 | 4.3 | 409 | 1 | ABA25687 | Probe #4153 for ge |
| c 248 | 324.7 | 4.3 | 409 | 1 | AAK29717 | Human bone marrow |
| c 249 | 324.7 | 4.3 | 409 | 1 | AAK04230 | Human brain expres |
| c 250 | 324.7 | 4.3 | 409 | 1 | ABS29354 | Human liver single |
| c 251 | 324.7 | 4.3 | 409 | 1 | AAI04133 | Probe #4124 used t |
| c 252 | 324.7 | 4.3 | 409 | 1 | ABS04271 | Human genome-deriv |

| | | | | | | |
|---|---|---|---|---|---|---|
| c 253 | 315.5 | 4.2 | 414 | 1 | AAI14380 | Probe #4313 for ge |
| c 254 | 315.5 | 4.2 | 414 | 1 | ABA56109 | Human foetal liver |
| c 255 | 315.5 | 4.2 | 414 | 1 | AAI35756 | Probe #4442 used t |
| c 256 | 315.5 | 4.2 | 414 | 1 | ABA45604 | Human breast cell |
| c 257 | 315.5 | 4.2 | 414 | 1 | ABA25761 | Probe #4227 for ge |
| c 258 | 315.5 | 4.2 | 414 | 1 | AAK29793 | Human bone marrow |
| c 259 | 315.5 | 4.2 | 414 | 1 | AAK04297 | Human brain expres |
| c 260 | 315.5 | 4.2 | 414 | 1 | ABS29438 | Human liver single |
| c 261 | 315.5 | 4.2 | 414 | 1 | AAI04204 | Probe #4195 used t |
| c 262 | 315.5 | 4.2 | 414 | 1 | ABS04353 | Human genome-deriv |
| 263 | 306.3 | 4.1 | 374 | 1 | AAV09123 | 5' nucleotide sequ |
| 264 | 306.3 | 4.1 | 374 | 1 | AAV15135 | Human adult testes |
| 265 | 300.7 | 4.0 | 415 | 1 | AEA20777 | Novel human polynu |
| 266 | 300.7 | 4.0 | 415 | 1 | AEA20798 | Novel human polynu |
| c 267 | 292.1 | 3.9 | 336 | 1 | ABA73835 | Human foetal liver |
| c 268 | 292.1 | 3.9 | 336 | 1 | AAI54279 | Probe #22965 used |
| c 269 | 292.1 | 3.9 | 336 | 1 | ABA38992 | Probe #17458 for g |
| c 270 | 292.1 | 3.9 | 336 | 1 | AAK48452 | Human bone marrow |
| c 271 | 292.1 | 3.9 | 336 | 1 | AAK22286 | Human brain expres |
| c 272 | 292.1 | 3.9 | 336 | 1 | ABS48146 | Human liver single |
| c 273 | 292.1 | 3.9 | 336 | 1 | ABS22263 | Human genome-deriv |
| c 274 | 285.7 | 3.8 | 387 | 1 | AAI27830 | Probe #17763 for g |
| c 275 | 285.7 | 3.8 | 387 | 1 | ABA76145 | Human foetal liver |
| c 276 | 285.7 | 3.8 | 387 | 1 | AAI56806 | Probe #25492 used |
| c 277 | 285.7 | 3.8 | 387 | 1 | AAK24809 | Human brain expres |
| c 278 | 285.7 | 3.8 | 387 | 1 | ABS50402 | Human liver single |
| c 279 | 285.7 | 3.8 | 387 | 1 | ABS24297 | Human genome-deriv |
| 280 | 283.3 | 3.8 | 341 | 1 | AAA98187 | Human retrovirus L |
| 281 | 270.3 | 3.6 | 343 | 1 | AAA98200 | Human retrovirus L |
| 282 | 269.9 | 3.6 | 343 | 1 | AAA98198 | Human retrovirus L |
| c 283 | 268.4 | 3.6 | 379 | 1 | AAI13385 | Probe #3318 for ge |
| c 284 | 268.4 | 3.6 | 379 | 1 | ABA55083 | Human foetal liver |
| c 285 | 268.4 | 3.6 | 379 | 1 | AAI34739 | Probe #3425 used t |
| c 286 | 268.4 | 3.6 | 379 | 1 | ABA44631 | Human breast cell |
| c 287 | 268.4 | 3.6 | 379 | 1 | AAK28798 | Human bone marrow |
| c 288 | 268.4 | 3.6 | 379 | 1 | AAK03348 | Human brain expres |
| c 289 | 268.4 | 3.6 | 379 | 1 | ABS28403 | Human liver single |
| c 290 | 268.4 | 3.6 | 379 | 1 | AAI03278 | Probe #3269 used t |
| c 291 | 268.4 | 3.6 | 379 | 1 | ABS03326 | Human genome-deriv |
| 292 | 265.1 | 3.5 | 343 | 1 | AAA98194 | Human retrovirus L |
| 293 | 262.9 | 3.5 | 341 | 1 | AAA98188 | Human retrovirus L |
| 294 | 262.7 | 3.5 | 310 | 1 | AAI21066 | Probe #10999 for g |
| 295 | 262.7 | 3.5 | 310 | 1 | ABA66142 | Human foetal liver |
| 296 | 262.7 | 3.5 | 310 | 1 | AAI46325 | Probe #15011 used |
| 297 | 262.7 | 3.5 | 310 | 1 | ABA48260 | Human breast cell |
| 298 | 262.7 | 3.5 | 310 | 1 | ABA33210 | Probe #11676 for g |
| 299 | 262.7 | 3.5 | 310 | 1 | AAK40307 | Human bone marrow |
| 300 | 262.7 | 3.5 | 310 | 1 | AAK14563 | Human brain expres |
| 301 | 262.7 | 3.5 | 310 | 1 | ABS39877 | Human liver single |
| 302 | 262.7 | 3.5 | 310 | 1 | AAI06786 | Probe #6777 used t |
| 303 | 262.7 | 3.5 | 310 | 1 | ABS14331 | Human genome-deriv |
| 304 | 260.3 | 3.5 | 343 | 1 | AAA98192 | Human retrovirus L |
| 305 | 260.3 | 3.5 | 343 | 1 | AAA98197 | Human retrovirus L |
| 306 | 258.7 | 3.5 | 343 | 1 | AAA98193 | Human retrovirus L |
| 307 | 255.5 | 3.4 | 343 | 1 | AAA98195 | Human retrovirus L |
| 308 | 255.5 | 3.4 | 343 | 1 | AAA98196 | Human retrovirus L |
| 309 | 253.9 | 3.4 | 343 | 1 | AAA98190 | Human retrovirus L |
| 310 | 252.3 | 3.4 | 343 | 1 | AAA98199 | Human retrovirus L |
| 311 | 251.2 | 3.4 | 323 | 1 | AAS84198 | DNA encoding novel |
| 312 | 250.7 | 3.3 | 343 | 1 | AAA98191 | Human retrovirus L |
| 313 | 247.3 | 3.3 | 333 | 1 | ACH83988 | Human genome deriv |
| 314 | 244.5 | 3.3 | 275 | 1 | AAI23208 | Probe #13141 for g |
| 315 | 244.5 | 3.3 | 275 | 1 | AAI48529 | Probe #17215 used |
| 316 | 244.5 | 3.3 | 275 | 1 | ABA50362 | Human breast cell |
| 317 | 244.5 | 3.3 | 275 | 1 | AAK42453 | Human bone marrow |
| 318 | 244.5 | 3.3 | 275 | 1 | ABS42065 | Human liver single |
| 319 | 238 | 3.2 | 279 | 1 | AAV09125 | 3' nucleotide sequ |
| 320 | 238 | 3.2 | 279 | 1 | AAV15136 | Human adult testes |
| 321 | 237.1 | 3.2 | 310 | 1 | AAV43214 | Multiple sclerosis |
| 322 | 237.1 | 3.2 | 310 | 1 | AAX29701 | Clone CL6-5' from |
| 323 | 237.1 | 3.2 | 310 | 1 | ADG14841 | MSRV associated cD |
| c 324 | 236 | 3.2 | 266 | 1 | ACH81073 | Human genome deriv |
| 325 | 234.4 | 3.1 | 322 | 1 | AAA98189 | Human retrovirus L |
| 326 | 233.1 | 3.1 | 310 | 1 | ADB84395 | MSRV-1 associated |
| c 327 | 229.1 | 3.1 | 321 | 1 | ADF59598 | Human polynucleoti |
| 328 | 215.2 | 2.9 | 297 | 1 | AAT02816 | Multiple sclerosis |
| 329 | 215.2 | 2.9 | 297 | 1 | AAT37974 | MSRV-1 POL gene cl |
| 330 | 215.2 | 2.9 | 297 | 1 | AAT96526 | MSRV-1B clone F1-1 |
| 331 | 215.2 | 2.9 | 297 | 1 | AAV43116 | Multiple sclerosis |
| 332 | 215.2 | 2.9 | 297 | 1 | ADB84550 | MSRV-1 DNA clone F |
| 333 | 215.2 | 2.9 | 297 | 1 | ADG14743 | MSRV-1 clone F-1-1 |
| 334 | 207.4 | 2.8 | 211 | 1 | ADD26898 | Human adipocyte Se |
| 335 | 205 | 2.7 | 278 | 1 | AAI20303 | Probe #10236 for g |
| 336 | 205 | 2.7 | 278 | 1 | ABA65340 | Human foetal liver |
| 337 | 205 | 2.7 | 278 | 1 | AAI45507 | Probe #14193 used |
| 338 | 205 | 2.7 | 278 | 1 | ABA47449 | Human breast cell |
| 339 | 205 | 2.7 | 278 | 1 | ABA32438 | Probe #10904 for g |
| 340 | 205 | 2.7 | 278 | 1 | AAK39493 | Human bone marrow |
| 341 | 205 | 2.7 | 278 | 1 | AAK13751 | Human brain expres |
| 342 | 205 | 2.7 | 278 | 1 | ABS39084 | Human liver single |
| 343 | 205 | 2.7 | 278 | 1 | AAI06006 | Probe #5997 used t |
| 344 | 205 | 2.7 | 278 | 1 | ABS13584 | Human genome-deriv |
| c 345 | 177.2 | 2.4 | 200 | 1 | AAI23581 | Probe #13514 for g |
| c 346 | 177.2 | 2.4 | 200 | 1 | ABA68694 | Human foetal liver |
| c 347 | 177.2 | 2.4 | 200 | 1 | AAI48896 | Probe #17582 used |
| c 348 | 177.2 | 2.4 | 200 | 1 | ABA50726 | Human breast cell |
| c 349 | 177.2 | 2.4 | 200 | 1 | ABA35659 | Probe #14125 for g |
| c 350 | 177.2 | 2.4 | 200 | 1 | AAK42820 | Human bone marrow |
| c 351 | 177.2 | 2.4 | 200 | 1 | AAK17036 | Human brain expres |
| c 352 | 177.2 | 2.4 | 200 | 1 | ABS42449 | Human liver single |
| c 353 | 177.2 | 2.4 | 200 | 1 | AAI09197 | Probe #9188 used t |
| c 354 | 177.2 | 2.4 | 200 | 1 | ABS16875 | Human genome-deriv |
| 355 | 156.6 | 2.1 | 202 | 1 | AAI56898 | Probe #25584 used |
| 356 | 156.2 | 2.1 | 207 | 1 | AAI46347 | Probe #15033 used |
| 357 | 156.2 | 2.1 | 207 | 1 | ABS39898 | Human liver single |
| 358 | 152.6 | 2.0 | 201 | 1 | AAK83421 | Human immune/haema |
| c 359 | 145.2 | 1.9 | 169 | 1 | ACH84817 | Human genome deriv |
| c 360 | 141.9 | 1.9 | 194 | 1 | ACH92789 | Human genome deriv |
| 361 | 139.2 | 1.9 | 150 | 1 | AAT96481 | Fragment POL2B of |
| 362 | 139.2 | 1.9 | 150 | 1 | AAV43153 | Multiple sclerosis |
| 363 | 139.2 | 1.9 | 150 | 1 | ADB84331 | MSRV-1 POL2B gene |
| 364 | 139.2 | 1.9 | 150 | 1 | ADG14777 | MSRV-1 pol cDNA. |
| 365 | 135.6 | 1.8 | 174 | 1 | ACH84800 | Human genome deriv |
| c 366 | 128.9 | 1.7 | 180 | 1 | ADF58618 | Human polynucleoti |
| c 367 | 127.3 | 1.7 | 180 | 1 | ADF58263 | Human polynucleoti |
| 368 | 122.6 | 1.6 | 152 | 1 | AAI23701 | Probe #13634 for g |
| 369 | 122.6 | 1.6 | 152 | 1 | ABA68815 | Human foetal liver |
| 370 | 122.6 | 1.6 | 152 | 1 | AAI49015 | Probe #17701 used |
| 371 | 122.6 | 1.6 | 152 | 1 | ABA50844 | Human breast cell |
| 372 | 122.6 | 1.6 | 152 | 1 | AAK42939 | Human bone marrow |
| 373 | 122.6 | 1.6 | 152 | 1 | ABS42569 | Human liver single |
| 374 | 122.6 | 1.6 | 152 | 1 | AAI09320 | Probe #9311 used t |
| 375 | 122.6 | 1.6 | 152 | 1 | ABS17000 | Human genome-deriv |
| 376 | 121.8 | 1.6 | 133 | 1 | AAI57411 | Probe #26097 used |
| 377 | 121 | 1.6 | 155 | 1 | AAI21346 | Probe #11279 for g |
| 378 | 121 | 1.6 | 155 | 1 | AAI46625 | Probe #15311 used |
| 379 | 121 | 1.6 | 155 | 1 | ABA33486 | Probe #11952 for g |
| 380 | 121 | 1.6 | 155 | 1 | AAK40578 | Human bone marrow |
| 381 | 121 | 1.6 | 155 | 1 | AAK14842 | Human brain expres |
| 382 | 121 | 1.6 | 155 | 1 | ABS40144 | Human liver single |
| 383 | 121 | 1.6 | 155 | 1 | ABS14555 | Human genome-deriv |
| c 384 | 101.8 | 1.4 | 135 | 1 | AAI28143 | Probe #18076 for g |
| c 385 | 101.8 | 1.4 | 135 | 1 | ABA76489 | Human foetal liver |
| c 386 | 101.8 | 1.4 | 135 | 1 | AAI57175 | Probe #25861 used |
| c 387 | 101.8 | 1.4 | 135 | 1 | ABA41003 | Probe #19469 for g |
| c 388 | 101.8 | 1.4 | 135 | 1 | AAK51142 | Human bone marrow |
| c 389 | 101.8 | 1.4 | 135 | 1 | AAK25134 | Human brain expres |
| c 390 | 101.8 | 1.4 | 135 | 1 | ABS50702 | Human liver single |
| c 391 | 101.8 | 1.4 | 135 | 1 | ABS24663 | Human genome-deriv |
| c 392 | 93 | 1.2 | 110 | 1 | ABA72820 | Human foetal liver |
| c 393 | 93 | 1.2 | 110 | 1 | AAI53246 | Probe #21932 used |
| c 394 | 93 | 1.2 | 110 | 1 | ABA38438 | Probe #16904 for g |
| c 395 | 93 | 1.2 | 110 | 1 | AAK47408 | Human bone marrow |
| c 396 | 93 | 1.2 | 110 | 1 | AAK21251 | Human brain expres |
| c 397 | 93 | 1.2 | 110 | 1 | ABS47149 | Human liver single |
| c 398 | 93 | 1.2 | 110 | 1 | ABS21546 | Human genome-deriv |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 399 | 92 | 1.2 | 126 | 1 | AAT37982 | Rheumatoid polyart |
| c | 400 | 89.2 | 1.2 | 103 | 1 | AAI22615 | Probe #12548 for g |
| c | 401 | 89.2 | 1.2 | 103 | 1 | ABA67692 | Human foetal liver |
| c | 402 | 89.2 | 1.2 | 103 | 1 | AAI47908 | Probe #16594 used |
| c | 403 | 89.2 | 1.2 | 103 | 1 | ABA49778 | Human breast cell |
| c | 404 | 89.2 | 1.2 | 103 | 1 | AAK41849 | Human bone marrow |
| c | 405 | 89.2 | 1.2 | 103 | 1 | AAK16108 | Human brain expres |
| c | 406 | 89.2 | 1.2 | 103 | 1 | ABS41446 | Human liver single |
| c | 407 | 89.2 | 1.2 | 103 | 1 | AAI08291 | Probe #8282 used t |
| c | 408 | 89.2 | 1.2 | 103 | 1 | ABS15874 | Human genome-deriv |
| | 409 | 85.2 | 1.1 | 107 | 1 | AAI20302 | Probe #10235 for g |
| | 410 | 85.2 | 1.1 | 107 | 1 | ABA65339 | Human foetal liver |
| | 411 | 85.2 | 1.1 | 107 | 1 | AAI45506 | Probe #14192 used |
| | 412 | 85.2 | 1.1 | 107 | 1 | ABA47448 | Human breast cell |
| | 413 | 85.2 | 1.1 | 107 | 1 | ABA32437 | Probe #10903 for g |
| | 414 | 85.2 | 1.1 | 107 | 1 | AAK39492 | Human bone marrow |
| | 415 | 85.2 | 1.1 | 107 | 1 | AAK13750 | Human brain expres |
| | 416 | 85.2 | 1.1 | 107 | 1 | ABS39083 | Human liver single |
| | 417 | 85.2 | 1.1 | 107 | 1 | AAI06005 | Probe #5996 used t |
| | 418 | 85.2 | 1.1 | 107 | 1 | ABS13583 | Human genome-deriv |
| c | 419 | 85 | 1.1 | 104 | 1 | ACH86125 | Human genome deriv |
| | 420 | 79.2 | 1.1 | 111 | 1 | AAT37979 | MSRV-1B POL gene f |
| | 421 | 79.2 | 1.1 | 111 | 1 | AAT96515 | MSRV-1B general co |
| | 422 | 79.2 | 1.1 | 111 | 1 | AAV43121 | Multiple sclerosis |
| | 423 | 79.2 | 1.1 | 111 | 1 | ADB84555 | MSRV-1B pol region |
| | 424 | 79.2 | 1.1 | 111 | 1 | ADG14748 | MSRV-1B consensus |

ALIGNMENTS

```
RESULT 1
AAX25665
ID   AAX25665 standard; cDNA to mRNA; 7582 BP.
XX
AC   AAX25665;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Complete human endogenous retrovirus W genome.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 71-74; 106pp; French.
XX
CC   This sequence represents the complete sequence of the human endogenous
CC   retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC   peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC   sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC   insulin- dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;

  Query Match             99.9%;  Score 7487.42;  DB 1;  Length 7582;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 7578;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
```

```
Qy      961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020

Qy     1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080

Qy     1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140

Qy     1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200

Qy     1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260

Qy     1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320

Qy     1321 GAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 GAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTG 1380

Qy     1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC 1440

Qy     1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
            ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAUGAATTACTCAATGATGTCCACCATA 1500

Qy     1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560

Qy     1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620

Qy     1621 TATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 TATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAG 1680

Qy     1681 CAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAG 1740
            ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     1681 CAAAACTTAGAAACCCTATTGAACTTGGCAACTTCGGTTTTTTATAATAGAGATCAGGAG 1740

Qy     1741 GAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 GAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCT 1800

Qy     1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAA 1860

Qy     1861 TAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAG 1920

Qy     1921 TAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 TAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980

Qy     1981 CAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTG 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 CAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTG 2040

Qy     2041 AGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTT 2100
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 AGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTT 2100

Qy     2101 GACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160

Qy     2161 CTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 CTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG 2220

Qy     2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280

Qy     2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340

Qy     2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400

Qy     2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460

Qy     2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520

Qy     2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580

Qy     2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640

Qy     2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGG 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGG 2700

Qy     2701 AGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATAC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 AGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATAC 2760

Qy     2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTG 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTG 2820

Qy     2821 TACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAAC 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAAC 2880

Qy     2881 TCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCAT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 TCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCAT 2940

Qy     2941 TAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATG 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 TAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATG 3000

Qy     3001 ATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMA 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 ATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMA 3060

Qy     3061 ATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGC 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 ATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGC 3120

Qy     3121 AGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      3121 AGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYA 3180

Qy      3181 TCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3181 TCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTG 3240

Qy      3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300

Qy      3301 WATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAA 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 WATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAA 3360

Qy      3361 GYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 GYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420

Qy      3421 TGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3421 TGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGA 3480

Qy      3481 GTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3481 GTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540

Qy      3541 GATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3541 GATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600

Qy      3601 GTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3601 GTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660

Qy      3661 GTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAA 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3661 GTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAA 3720

Qy      3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAAC 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAAC 3780

Qy      3781 CCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATT 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3781 CCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATT 3840

Qy      3841 TCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3841 TCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTT 3900

Qy      3901 CAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3901 CAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960

Qy      3961 AGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYA 4020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3961 AGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYA 4020

Qy      4021 GCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAA 4080
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      4021 GCTRGCAGAACTAATAGCCYTCAYTNGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAA 4080

Qy      4081 TATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4081 TATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140

Qy      4141 AGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4141 AGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200

Qy      4201 ATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCA 4260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4201 ATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCA 4260

Qy      4261 NAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGC 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4261 NAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGC 4320

Qy      4321 TGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCC 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4321 TGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCC 4380

Qy      4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG 4440

Qy      4441 CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAAC 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4441 CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAAC 4500

Qy      4501 TATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCAC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4501 TATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCAC 4560

Qy      4561 CCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARAT 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4561 CCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARAT 4620

Qy      4621 AKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTC 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4621 AKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTC 4680

Qy      4681 AGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4681 AGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740

Qy      4741 CCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4741 CCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCA 4800

Qy      4801 RAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAA 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4801 RAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAA 4860

Qy      4861 ATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4861 ATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920

Qy      4921 GCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4921 GCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980

Qy      4981 TGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTA 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4981 TGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTA 5040

Qy      5041 AAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAA 5100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5041 AAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAA 5100

Qy      5101 AGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5101 AGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160

Qy      5161 CCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCA 5220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5161 CCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGACATCA 5220

Qy      5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAG 5280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAG 5280

Qy      5281 AGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5281 AGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340
```

```
Qy      5341 CCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTG 5400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5341 CCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTG 5400

Qy      5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATAC 5460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATAC 5460

Qy      5461 TGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGG 5520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5461 TGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGG 5520

Qy      5521 ATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCC 5580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5521 ATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCC 5580

Qy      5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640

Qy      5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700

Qy      5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760

Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTCTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020

Qy      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080

Qy      7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140

Qy      7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGC 7200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGC 7200

Qy      7201 GGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGA 7260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7201 GGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGA 7260

Qy      7261 GACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7261 GACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320

Qy      7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380

Qy      7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440

Qy      7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500

Qy      7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||||||||||
Db       7561 TCTATTAAATCTTGCARCTGCR 7582


RESULT 2
AAS76475
ID   AAS76475 standard; cDNA; 4349 BP.
XX
AC   AAS76475;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12279.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12288.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 12279; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 4349 BP; 1319 A; 1061 C; 941 G; 1026 T; 0 U; 2 Other;

  Query Match             47.4%;  Score 3549.5;  DB 1;  Length 4349;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches 3974;  Conservative  148;  Mismatches  195;  Indels  237;  Gaps   28;
```

```
Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGA---RGCGCCCATTGCCGCTCCCAATTGGGCTA 291
              |||||||||||||||||||||||||||||   :|||||||||||| |||||| | ||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGGCGCCCATTGCCACTCCCAGTCGGGCTA 180

Qy        292 AAGGCTTGCCATTGTNCCTGCAC-GGCTAAGTGCCT-GGGTTTGTTCTAATTGAGCTGAA 349
              |||||  ||||||| ||||||| |||||||||||| |||||  | ||||| ||||||||
Db        181 GAGGCTCACCATTGTTCCTGCACAGGCTAAGTGCCTGGGGTTCATCCTAATCGAGCTGAA 240

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| |||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        241 CACTAATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAA 300

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              ||||||||||||||||||||||||||||||||||||||||||||:||| ||||||||||
Db        301 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGG 359

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGT 529
              ||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||||
Db        360 TCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGT 419

Qy        530 TCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMS 589
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||::
Db        420 TCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAAC 479

Qy        590 RACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTA 649
              :||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTA 539

Qy        650 TTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCT 709
              |||||||:||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        540 TTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCT 599

Qy        710 TCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTA 769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTA 659

Qy        770 GACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCA 829
              |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCA 719

Qy        830 TTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 TTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAG 779

Qy        890 AGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTC 949
              ||||||||||||||||||||||||||||||||||||||::||||||||||||||||||||
Db        780 AGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTC 839

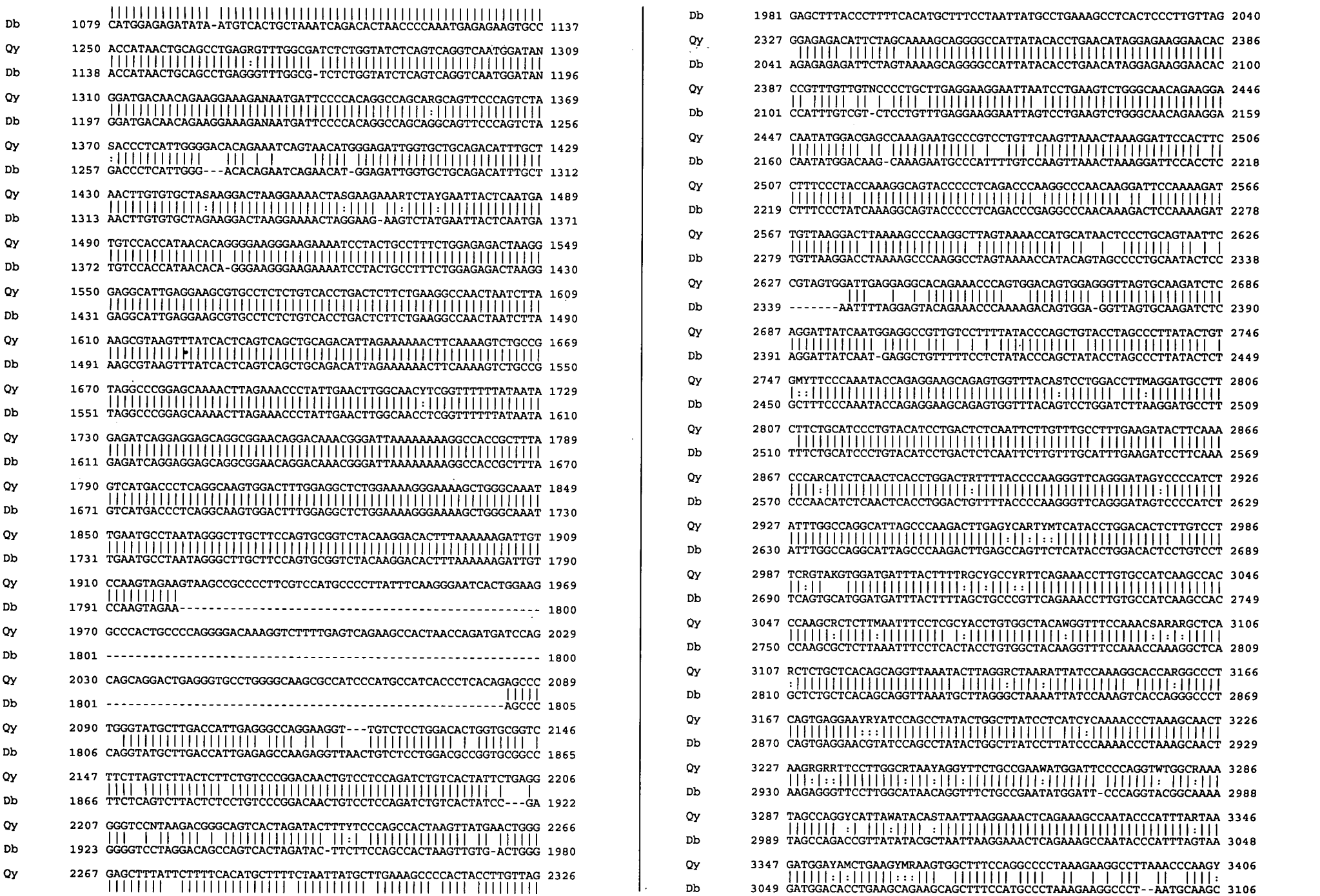
Qy        950 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAA 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAA 899

Qy       1010 TATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAG 1069
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        900 TATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAG 958

Qy       1070 TGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCT 1129
              |||||||||:||||::|||||||||||||||||||||||||||||||||||||||||||||
Db        959 TGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCT 1018

Qy       1130 CAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGA 1189
              |||||||:|||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1019 CAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGA 1078

Qy       1190 CATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCC 1249
```



```
Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCAC------------ 3454
             ||||||| ||||:|||| |||:|||||||||||||: |:|||::||||
Db       3107 CCCAGTGCTAAGCTTGCTAACGGGGCAAGACTTTTCCTCATATGTCACAAAAAACAAACA 3166

Qy       3455 ---------AGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
                     | ||||||||| :||:|||||||: |||||||| | | |:|| :|||||:||
Db       3167 AACAAAGAAAAAAAAAAACATGAATAGCTCTAAGAGTCCTTGCGCTGGTCTGAGGGACGA 3226

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
             | |||||||::||||::|||||| ||||||||:|||||||||||||||||||  :|:||
Db       3227 GTTTGCAACCCGTGGCATACCTGA-TAAGGAAATTGATGTAGTGGCAAAGGGTCAGCCTC 3285

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
             ||||||||:|||||||||||||||||||||||::|||||||||||||||||||||||||||
Db       3286 ATTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATAC 3345

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
             ||||:||| |||||||||||||||||||||||||:|||||::|||||||||||||||||||
Db       3346 AGGGAAGAAATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAG 3405

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
             |||||||||||||||||||||||: ||||||||||:|||||||||||||||||||:|||
Db       3406 GAGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGGC 3465

Qy       3746 CAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATG 3805
             || ||||||:|||||||||||||||| |||||||||  | | |||||||||||| ||||||
Db       3466 CAATGCTGCGACTGTGCACTTGTGCAGCTCTTAACCTGGACACATTTCTTCCACACAATG 3525

Qy       3806 AAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
             ||||||||||:||:|||||||||||||||:||||| ||||||||||||||||||||||||
Db       3526 AAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCCACTCGAGGGG 3585

Qy       3866 ACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCC 3925
             ||| |:||||:|||||:|||||||||||: ||| | |||||||||||||||||||| |||
Db       3586 ACCATTTAGAGGTTCCCTTGACTGATCCCAACCCT-AACTTGTATACTGATGGAAGATCC 3644

Qy       3926 TTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTT 3985
             |||||||||||||||||| |||||:|||||||||||| ||||||||||||||||||:||
Db       3645 TTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTAGTCAGTGATAATGGAATACTT 3704

Qy       3986 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYT 4045
             ||||||||||||||||||||||||||||||||:||||:||| |||         |||:|
Db       3705 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAAC---------TCACT 3755

Qy       4046 KGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCT 4105
              |||||||||||||||||||||:|||||||:|||||||||||||||||||:|:||||||
Db       3756 CGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCT 3815

Qy       4106 YACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAG 4165
              :|||||||| |||||||||||::| |||||:|:|||||||||||||||||||||:|||
Db       3816 TACCTAGTCCTCCATGCCCATGCAGTAATATGGAGAGAAAGGGAATTCCTAACTTCTGAG 3875

Qy       4166 RGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCT 4225
              :|||||||||||:|||||||||||||||||| :|||||||:|||| ||||||||||||
Db       3876 GGAACACCTATCAAACATCAGGAAGCCATTAGGGAATTATTATTGGCCGTACAGAAACCT 3935

Qy       4226 ARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAA 4285
              |:|||||||:|||||||||||||:|||||||||| |||||||||:|||||||||||:||
Db       3936 AAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAA 3995

Qy       4286 GRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAA 4345
              :|||: |||||||:|||||||||:|||||||||||||||||||||||||||||||||||
Db       3996 AGGAACCGCCAAGCAGATATTGAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAA 4055

Qy       4346 ATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTAC 4405
             ||||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||
Db       4056 ATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTAT 4115

Qy       4406 TCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGG 4462
             |||| || ||||| |||||||||| |||||||||    |||||||||||||| ||| ||
Db       4116 TCAGAAGAAGAAATAGAATGGGGAATCTCACGAGGACATAGTTTTCTCCCCTCAGGATGG 4175

Qy       4463 TTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAA 4522
              |||||| ||| |||| ||||||| |||||||||| ||| ||||||||||||||||||
Db       4176 CAAGCCACCGAACAAGGAAAAATACCTTTGCCTGCAGCTAACCAATGGAAATTACTTAAA 4235

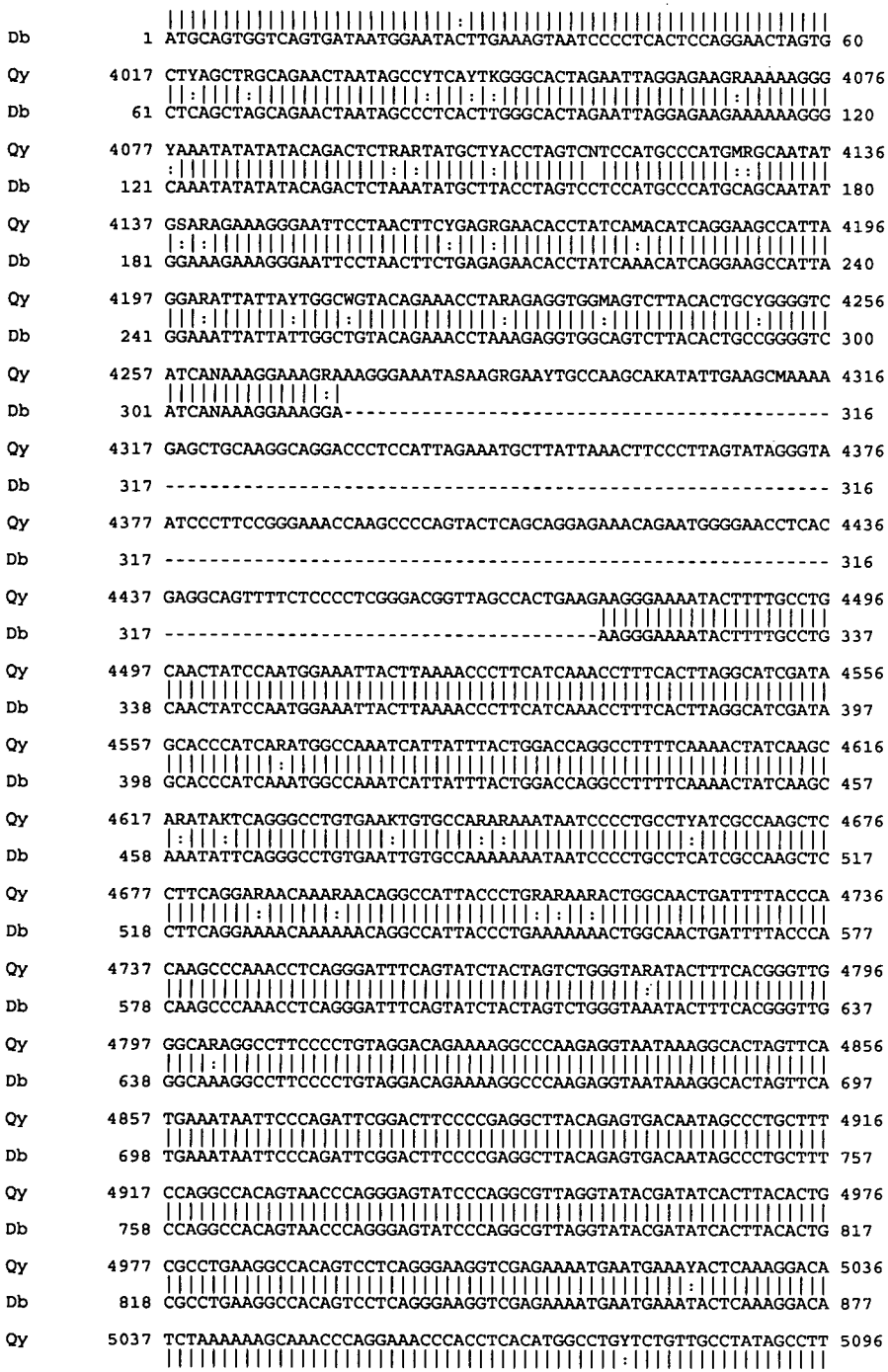
Qy       4523 ACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTA 4582
              |||  ||| |||| |||||||||| |||| |||||||||||||:|||||||||  |||
Db       4236 ACCACTCACCAAAACTTTCACTTATGCATTGATAGCACCCATCAGATGGCCAAATTGTTA 4295

Qy       4583 TTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAA 4636
              |||||||||||||| |||||||||||| ||||||:|||:|  |||||||| ||
Db       4296 TTTACTGGACCAGGACTTTTCAAAACTCTCAAGCAGATAGTTGGGGCCTGTAAA 4349
```

RESULT 3
ADE09587
ID  ADE09587 standard; DNA; 3464 BP.
XX
AC  ADE09587;
XX
DT  29-JAN-2004  (first entry)
XX
DE  Novel DNA-related contig nucleotide sequence #309.
XX
KW  novel gene; novel protein; tissue marker; molecular weight marker;
KW  chromosome marker; genetic disorder; contig; ds.
XX
OS  Unidentified.
XX
PN  WO2003054152-A2.
XX
PD  03-JUL-2003.
XX
PF  10-DEC-2002; 2002WO-US039555.
XX
PR  10-DEC-2001; 2001US-0339739P.
PR  11-DEC-2001; 2001US-0339453P.
PR  14-MAR-2002; 2002US-0365091P.
PR  14-MAR-2002; 2002US-0365384P.
PR  12-APR-2002; 2002US-0372381P.
PR  12-APR-2002; 2002US-0372615P.
PR  22-APR-2002; 2002US-00128558.
PR  24-APR-2002; 2002US-0376045P.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI  Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI  Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR  WPI; 2003-569235/53.
XX
PT  New polynucleotides, useful for expressing recombinant proteins for
PT  analysis, characterization or therapeutic use, or as markers for tissues
PT  in which the corresponding protein is preferentially expressed.
XX
PS  Disclosure; SEQ ID NO 2131; 1177pp; English.
XX
CC  The invention comprises the amino acid and coding sequences of novel
CC  proteins. The DNA and protein sequences of the invention are useful as:
CC  markers for tissues in which the corresponding protein is preferentially
CC  expressed; as molecular weight markers on gels; as chromosome markers or
CC  tags; to identify chromosomes or to map related gene positions; and to
CC  compare with endogenous DNA sequences in patients to identify potential
CC  genetic disorders. The present DNA sequence was used in the
CC  exemplification of the invention.
XX
SQ  Sequence 3464 BP; 1038 A; 913 C; 666 G; 845 T; 0 U; 2 Other;

  Query Match             40.4%;  Score 3025.7;  DB 1;  Length 3464;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches 3252;  Conservative   78;  Mismatches   93;  Indels  235;  Gaps   10;

```
Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
```


Db 1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60
Qy 4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db 61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120
Qy 4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db 121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180
Qy 4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db 181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240
Qy 4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db 241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300
Qy 4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
Db 301 ATCANAAAGGAAAGGA---------- 316
Qy 4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db 317 ---------- 316
Qy 4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db 317 ---------- 316
Qy 4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
Db 317 ----------AAGGGAAAATACTTTTGCCTG 337
Qy 4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
Db 338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397
Qy 4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
Db 398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457
Qy 4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
Db 458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517
Qy 4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
Db 518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577
Qy 4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
Db 578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637
Qy 4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
Db 638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697
Qy 4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
Db 698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757
Qy 4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
Db 758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817
Qy 4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
Db 818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877
Qy 5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
Db 878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937
Qy 5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
Db 938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997
Qy 5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
Db 998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057
Qy 5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
Db 1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117
Qy 5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
Db 1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177
Qy 5337 AATTCCCCATCCCTA-GATACATCCTGGG----AAGGACCCTACCCAGT---CATTTTAT 5388
Db 1178 AATTCCCCATCCCTAGGATACATCCTGGGGGAAGGGACCCTTACCCAGTTCATTTTTATT 1237
Qy 5389 YTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 5448
Db 1238 TTACCCCAATTGCGGGTTTAAAGTGGCTGGAGTGGTCTTGGATACATCACACTTGAGTCA 1297
Qy 5449 AATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 5508
Db 1298 AATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 1357
Qy 5509 AACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAAT 5568
Db 1358 AACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAAT 1417
Qy 5569 CATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTST 5628
Db 1418 CATAAATCCCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCT 1477
Qy 5629 TTCACTCTCACTGCACCCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAG 5688
Db 1478 TTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAG 1537
Qy 5689 TTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTST 5748
Db 1538 TTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCT 1597
Qy 5749 AAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCC 5808
Db 1598 AAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCC 1657
Qy 5809 ACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGT 5868
Db 1658 ACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGT 1717
Qy 5869 CCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGAT 5928
Db 1718 CCTGGAGGACTTGGAGTCACTGTCTGTCGGACTTACTTCACCCAAACTGGTATGTCTGAT 1777
Qy 5929 GGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTC 5988
Db 1778 GGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGCAATCTCCCAACTG 1837
Qy 5989 ACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAA 6048
Db 1838 ACCCGGGGACATAGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAA 1897
Qy 6049 ACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCAT 6108
Db 1898 ACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCAT 1957
Qy 6109 GAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
Db 1958 GAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCA 2017

```
Qy       6169 TATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACT 6228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2018 TATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACT 2077

Qy       6229 TCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACC 6288
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2078 TCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACC 2137

Qy       6289 TGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACT 6348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2138 TGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACT 2197

Qy       6349 CCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCC 6408
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2198 CCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCC 2257

Qy       6409 TATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCY 6468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db       2258 TATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCT 2317

Qy       6469 ATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCA----TATCTAAGCCCCGCAA 6524
             |||:|||||||||||||||||||||||||||||||||||||      |   |  ||||||||
Db       2318 ATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAATATCCTAAAGCCCCCGCAA 2377

Qy       6525 CAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTAC 6584
             |||||||||||||||||||||||||||||| |||||||||| ||||| | |||||||||
Db       2378 CAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTACTAGGCGGACTAGGTAC 2437

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             ||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db       2438 TGGCATTGGCGGTATTACAACCTCTACTCAGTTCTACCACAAACTATCTCAAGAACTAAA 2497

Qy       6645 TGGGGACATGGAACGGGTCGCC--GACTCCCTGGTCACCTTGCAAGATCAACTTAACTCC 6702
             ||| |||||||||| |||||||      ||||||||||||||||||||||||||||||||
Db       2498 TGGTGACATGGAACAGGTCGCCCGACTCCCCTGGTCACCTTGCAAGATCAACTTAACTCC 2557

Qy       6703 CTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGG 6762
             ||||||||||||||||||:||||||||||||||||||:||||||||||||:||||||
Db       2558 CTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGG 2617

Qy       6763 GGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTC 6822
             |||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Db       2618 GGAACCTGTTTACTTTTAGGGGAAGAATGCTGTTACTATGTTAATCAATCCGGAATCGTC 2677

Qy       6823 ACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAAC 6882
             ||||||||||||:|||||||||:||||||||||||||||:||||||:||||||||||||
Db       2678 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAAC 2737

Qy       6883 ACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCT 6942
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2738 ACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCT 2797

Qy       6943 CTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAAC 7002
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       2798 CTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAC 2857

Qy       7003 TTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAG 7062
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2858 TTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAG 2917

Qy       7063 ACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       2918 ACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGAC 2977

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2978 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 3037

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       3038 GCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 3097

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTT----------CCTAGGCTGAYT 7291
             ||||||||||||||||||||||||||||||||||||||          |||| |    :
Db       3098 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACCTAAGGANTCC 3157

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG---CTT 7348
                 || ||:||      |  |||  |   ||  |  ||   || ||||||||||    ||
Db       3158 CTAAAGCCTTAGCTGGGGAAGGGTGGACCACATCCACCTTTTTAAAACACGGGGGGCTTT 3217

Qy       7349 GCAACTTAGYTCACA---CCTGACCAATCAGAGAGCTCACTAAAATG-CTAATTAGGCAA 7404
             |||||||      ||   | |||||||||||||||||||||||||| ||||||||||||
Db       3218 GCAACTTTAGGTTCACACCTTGACCAATCAGAGAGCTCACTAAAATGCCTAATTAGGCAA 3277

Qy       7405 AGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAA 7464
             ||||||||||||||||||||||||||||||:|||||||:|||||||||||||||||||||
Db       3278 AGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAA 3337

Qy       7465 TGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCCTTTGGGTCCCCTC 7524
             |||||||||||||||||||||:|||||  ||||||||||||||||||||||||||||||||
Db       3338 TGATCGGGATATAAACCCAAGTCTTCGAAGCGGCAACGGCAACCCCCTTTGGGTCCCCTC 3397

Qy       7525 CCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||:
Db       3398 CCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 3455
```

RESULT 4
AAD24195
ID AAD24195 standard; cDNA; 2930 BP.
XX
AC AAD24195;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human syncytin cDNA.
XX
KW Human; syncytin; preeclampsia; gestational trophoblast disorder;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 930. .2546
FT /*tag= a
FT /product= "Syncytin"
XX
PN WO200204678-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US021719.
XX
PR 07-JUL-2000; 2000US-0216657P.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Keith JC, Mccoy JM, Mi S;
XX
DR WPI; 2002-171727/22.
DR P-PSDB; AAE14540.
XX
PT Identifying a compound for treating a subject with or at risk of
PT developing preeclampsia, comprises determining whether the expression or
PT activity of syncytin in the cell is modulated in the presence of a test
PT compound.
XX
PS Disclosure; Page 39-42; 43pp; English.
XX
CC The invention relates to identifying compounds which are modulators of
CC syncytin expression. The syncytin modulators are useful in diagnosis and
CC treatment of preeclampsia and gestational trophoblast disorders (e.g.
CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the envelope

```
CC   gene of human endogenous defective retrovirus, HERV-W. The present
CC   invention is based partly on the discovery that syncytin expression is
CC   dramatically reduced in preeclampsia, and is also mis-localised to the
CC   apical syncytiotrophoblast membrane. The present sequence is human
CC   syncytin cDNA
XX
SQ   Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

  Query Match            38.6%;  Score 2890.1;  DB 1;  Length 2930;
  Best Local Similarity  98.3%;  Pred. No. 0;
  Matches 2873;  Conservative   48;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:||||||||||||||:|:||:||
Db          8 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 67

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 127

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        128 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 187

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 247

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 307

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 367

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 427

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 487

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 547

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 607

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             ||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||
Db        608 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 667

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 727

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 787

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 847

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        848 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 907

Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
             ||||||||||||||||| |||||||:|||||||||||||||||||||:||:||||:
Db        908 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 966

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1026

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1086

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1146

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1206

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1266

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1326

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1386

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1446

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1506

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1566

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1567 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1626

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1686

Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1746

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1806

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
             ||||||||||:|||:|||||||||||||||||||||||||||| ||||||||||||||||
Db       1807 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1866

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1926

Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1986

Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1987 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2046

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
              |||||||||||||||||||||||:|||||||||||||||||||:||||||||||||:|
Db       2047 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2106

Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2166

Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
              |||||||||||||||:|||||||||:||||||||||||||||:||||||:|||||||
Db       2167 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2226

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2227 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2286

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2287 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2346

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2347 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2406

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
              |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       2407 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2466

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2526

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
              |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2527 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2586

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2587 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2646

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
              ||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       2647 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2706

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
              :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2707 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2766

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
              |||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db       2767 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2826

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
              |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2886

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||||||||||||||||||||||||||||:||||
Db       2887 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 2930


RESULT 5
AAX77526
ID   AAX77526 standard; cDNA; 2946 BP.
XX
AC   AAX77526;
XX
DT   10-AUG-1999  (first entry)
XX
DE   Human secreted protein AJ172_2 cDNA.
XX
KW   Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW   bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW   cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW   haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW   cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
OS   Homo sapiens.
XX
PN   WO9926972-A1.
XX
PD   03-JUN-1999.
XX
PF   17-NOV-1998;   98WO-US024614.
XX
PR   21-NOV-1997;   97US-00976110.
PR   18-MAY-1998;   98US-00080478.
PR   20-OCT-1998;   98US-00175928.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR   WPI; 1999-357813/30.
DR   P-PSDB; AAY08622.
XX
PT   New polynucleotides encoding secreted proteins.
XX
PS   Claim 13a; Page 100-101; 142pp; English.
XX
CC   This invention describes novel human secreted proteins encoded by
CC   polynucleotides isolated from human adult testes, adult brain, adult
CC   blood or adult placenta, or murine adult bone marrow or thymus cDNA
CC   libraries. The products of the invention are predicted to have biological
CC   activities which would make them suitable for treating, preventing or
CC   ameliorating medical conditions in humans and animals, although no
CC   supporting data is given. Suggested activities include nutritional
CC   activity, cytokine and cell proliferation/differentiation activity,
CC   immune stimulating (e.g. as vaccines) or suppressing activity,
CC   haematopoiesis regulating activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, cadherin/tumour invasion suppressor activity, and tumour
CC   inhibition activity. The polynucleotides are also stated to be useful for
CC   gene therapy
XX
SQ   Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              ||:|||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305
```

```
Qy      4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy      5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy      5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy      5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

Qy      5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy      5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             |||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||||
Db       606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy      5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy      5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785

Qy      5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845

Qy      5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905

Qy      5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
             |||||||||||||||||| |||||||:|||||||||||||||||||||:||:|||||:
Db       906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964

Qy      5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024

Qy      5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084

Qy      5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144

Qy      5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204

Qy      5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264

Qy      5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324

Qy      5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384

Qy      6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444

Qy      6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504

Qy      6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564

Qy      6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db      1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624

Qy      6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684

Qy      6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744

Qy      6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804

Qy      6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
             ||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864

Qy      6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924

Qy      6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984

Qy      6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044

Qy      6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
             ||||||||||||||||||||||||:|||||||||||||||||||:|||||||||||||:|
Db      2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104

Qy      6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164

Qy      6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
             ||||||||||||||||||:|||||||||:|||||||||||||||:|||||||:|||||||
Db      2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224

Qy      6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db      2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284

Qy      6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db      2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344

Qy      6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db      2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404

Qy      7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db      2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464

Qy      7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
              ||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
              :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
              |||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db       2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
              |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||||||||||||||||||||||||||||:||||:
Db       2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929


RESULT 6
AAZ59468
ID   AAZ59468 standard; cDNA; 2946 BP.
XX
AC   AAZ59468;
XX
DT   11-APR-2000  (first entry)
XX
DE   Human secreted protein AJ172_2 polynucleotide sequence.
XX
KW   Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW   placental pathology; metastasis inhibition; nutritional activity;
KW   immune stimulator; haematopoiesis regulator; tissue growth;
KW   tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9960020-A1.
XX
PD   25-NOV-1999.
XX
PF   17-MAY-1999;   99WO-US010915.
XX
PR   18-MAY-1998;   98US-00080478.
PR   20-OCT-1998;   98US-00175928.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Mi S,  Treacy M;
XX
DR   WPI; 2000-116311/10.
DR   P-PSDB; AAY67313.
XX
PT   New polynucleotides encoding secreted cDNA libraries, used to develop
PT   products for the diagnosis and treatment of neoplastic disease.
XX
PS   Claim 14; Page 107-108; 149pp; English.
XX
CC   This is the human secreted protein AJ172_2 nucleotide sequence, obtained
CC   from a human adult testes cDNA library. The invention relates to secreted
CC   human and murine proteins. The polynucleotides and proteins are predicted
CC   to have biological activities which would make them suitable for
CC   treating, preventing or ameliorating medical conditions in humans and
CC   animals. Detection of the levels of the proteins can be used for the
CC   diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC   which modulate the expression or function of the proteins may be used for
CC   treating a neoplastic disease and inhibiting metastasis. Other suggested
CC   activities include nutritional activity (e.g. in feeds), cytokine and
CC   cell proliferation/differentiation activity, immune stimulating (e.g. as
CC   vaccines) or suppressing activity, haematopoiesis regulating activity,
CC   tissue growth activity, activin/inhibin activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC   receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC   invasion suppressor activity, and tumour inhibition activity. The
CC   polynucleotide sequences are also stated to be useful for gene therapy
XX
SQ   Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              ||:|||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

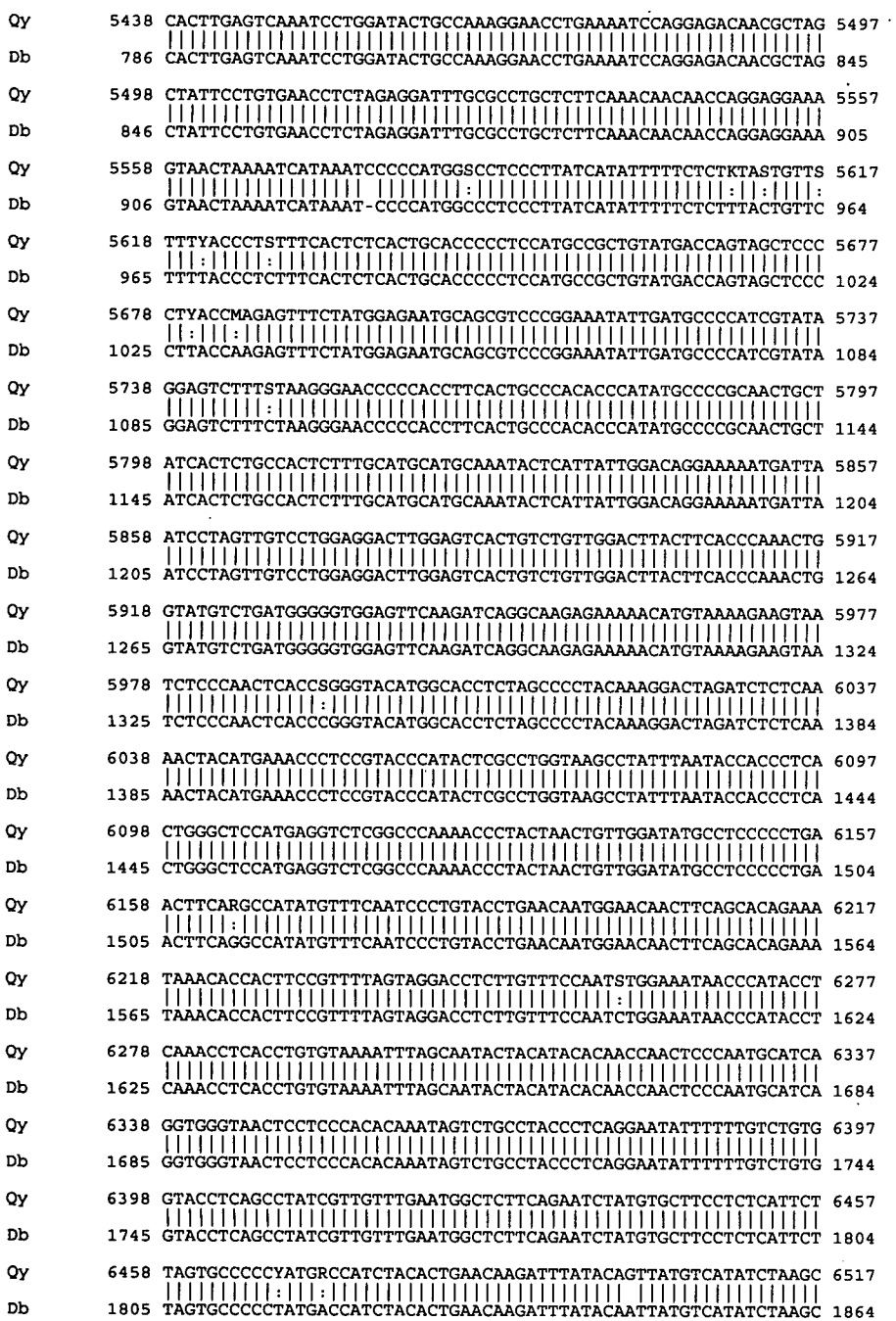
Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
              ||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db         606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785
```


Qy 5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
Db 786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845
Qy 5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
Db 846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905
Qy 5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
Db 906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964
Qy 5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
Db 965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024
Qy 5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
Db 1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084
Qy 5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
Db 1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144
Qy 5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
Db 1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204
Qy 5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
Db 1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264
Qy 5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
Db 1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324
Qy 5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
Db 1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384
Qy 6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
Db 1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444
Qy 6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
Db 1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504
Qy 6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
Db 1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564
Qy 6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
Db 1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624
Qy 6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
Db 1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684
Qy 6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
Db 1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744
Qy 6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
Db 1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804
Qy 6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
Db 1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864
Qy 6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
Db 1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924
Qy 6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
Db 1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984
Qy 6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
Db 1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044
Qy 6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
Db 2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104
Qy 6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
Db 2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164
Qy 6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
Db 2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224
Qy 6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
Db 2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284
Qy 6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
Db 2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344
Qy 6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
Db 2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404
Qy 7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
Db 2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464
Qy 7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
Db 2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524
Qy 7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
Db 2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584
Qy 7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
Db 2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644
Qy 7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
Db 2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704
Qy 7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
Db 2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764
Qy 7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
Db 2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824
Qy 7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
Db 2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884
Qy 7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
Db 2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929

```
RESULT 7
ADC38776
ID   ADC38776 standard; cDNA; 2946 BP.
XX
AC   ADC38776;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human cDNA encoding a secreted protein #63.
XX
KW   ss; gene; immune disorder; severe combined immunodeficiency; SCID;
KW   autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW   rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
KW   lymphoid cell deficiency; osteoporosis; osteoarthritis;
KW   peripheral nervous system disease; peripheral neuropathy;
KW   Alzheimer's disease; Parkinson's disease; coagulation disorder;
KW   inflammatory disease; systemic inflammatory response syndrome; SIRS;
KW   ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;
KW   hypersensitivity; regeneration; neural cell proliferation; fertility;
KW   tumour; chemokine; human; secreted protein.
XX
OS   Homo sapiens.
XX
PN   US2002193567-A1.
XX
PD   19-DEC-2002.
XX
PF   02-APR-2002; 2002US-00114893.
XX
PR   11-AUG-1995;   95US-00514014.
PR   05-APR-1996;   96US-00628364.
PR   19-APR-1996;   96US-00635311.
PR   07-JUN-1996;   96US-00659224.
PR   17-JUN-1996;   96US-00664596.
PR   09-JUL-1996;   96US-00677231.
PR   26-JUL-1996;   96US-00686878.
PR   23-AUG-1996;   96US-00701819.
PR   27-SEP-1996;   96US-00721488.
PR   27-SEP-1996;   96US-00721798.
PR   27-SEP-1996;   96US-00721923.
PR   27-SEP-1996;   96US-00721926.
PR   25-OCT-1996;   96US-00738367.
PR   30-OCT-1996;   96US-00739775.
PR   13-JAN-1997;   97US-00783395.
PR   10-APR-1997;   97US-00833823.
PR   02-JUN-1997;   97US-00867677.
PR   05-SEP-1997;   97US-00924838.
PR   06-OCT-1999;   99US-00413232.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M,  Bowman MR,  Spaulding V,  Carlin-Duckett M;
PI   Kelleher K;
XX
DR   WPI; 2003-657236/62.
DR   P-PSDB; ADC38777.
XX
PT   Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
PT   BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
PT   useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 134; 412pp; English.
XX
CC   The invention relates to a protein comprising fully defined AZ302 1
CC   protein or BD127 1 6 protein. The polynucleotides are useful for
CC   expressing recombinant proteins for analysis and are also useful as
CC   chromosome markers or tags to identify chromosomes or to map related gene
CC   positions. The proteins are useful as amino acid supplement, carbon
CC   source, nitrogen source and carbohydrate source. The proteins are useful
CC   for treating various immune deficiencies and disorders (e.g. severe
CC   combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC   sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
CC   reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
CC   osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
CC   peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
CC   coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
CC   response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),
CC   anaphylaxis and hypersensitivity. Proteins are also useful for inducing
CC   tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
CC   nerve growth or regeneration, for proliferating neural cells and for
CC   regenerating nerve and brain tissue, for inducing fertility and for
CC   inhibiting tumour growth. Proteins are also useful as chemokine for
CC   mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
CC   useful as inhibitors of receptor/ligand interactions. The present
CC   sequence represents cDNA encoding a human secreted protein.
XX
SQ   Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:|||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              ||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

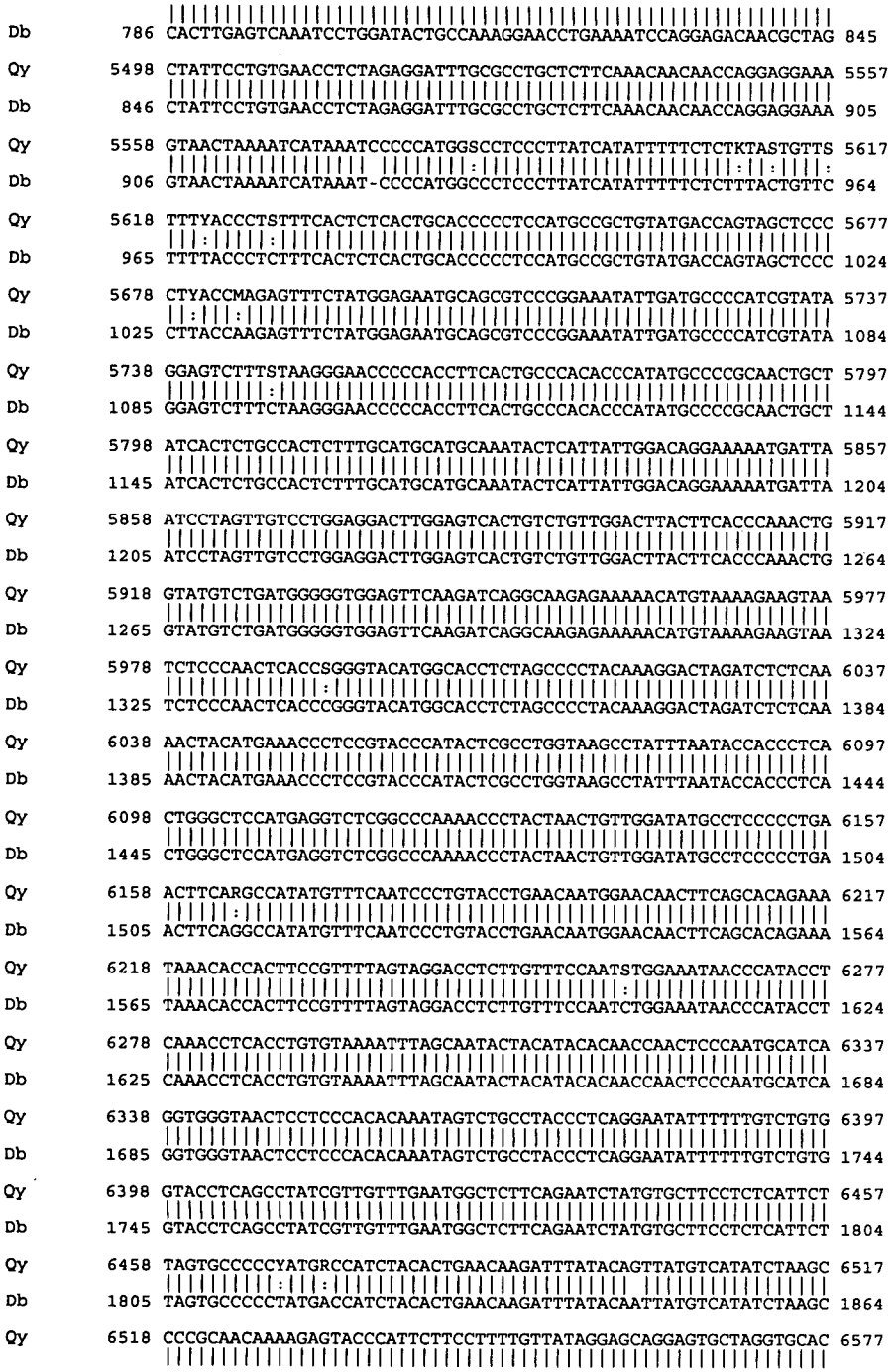
Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db         606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
```



RESULT 8
AAS71727

```
ID   AAS71727 standard; cDNA; 3831 BP.
XX
AC   AAS71727;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #7531.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG07540.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 7531; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3831 BP; 1173 A; 953 C; 815 G; 890 T; 0 U; 0 Other;

  Query Match           36.7%;  Score 2750.9;  DB 1;  Length 3831;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches 2814;  Conservative  140;  Mismatches   54;  Indels   33;  Gaps   10;

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||  |||||||||||||| ||||   ||||| |||| ||| || |||| ||||||||||||
Db        259 CATCGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCAAAAAAATTGTTAAGGAC 318

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
              ||||||||||||||| ||||||  |||||||  |  |||||| || | |       |
Db        319 CTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCAATACTCC-------AAC 371

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCA 2696
              |   |  | ||||||||||  |||||||| |||||||||||||||||||||||||||||
Db        372 TTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCA 430

Qy       2697 ATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAA 2756
              || ||||| ||||||| | |||||||||||||||||| |||||| | | ||::|||||||
Db        431 AT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTACATTCTGCTTTCCCAA 489

Qy       2757 ATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATC 2816
              |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||||
Db        490 ATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATC 549

Qy       2817 CCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCT 2876
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||:  |||
Db        550 CCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCT 609

Qy       2877 CAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAG 2936
              ||||||||||||||:|||||||||||||||||||||||:| |||||||||||||||||
Db        610 CAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCTCCATCTATTTGGCCAG 669

Qy       2937 GCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTG 2996
              |||||||||||||||||||:||:|::|||||||||||||||||||||||:|||: ||
Db        670 GCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCAGTATATG 729

Qy       2997 GATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTC 3056
              |||||||||||||||:|    ||::|||||||||||||||||||||||||||||||:|||
Db        730 GATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTC 789

Qy       3057 TTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCA 3116
              ||:||||||||||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||
Db        790 TTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCA 849

Qy       3117 CAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAA 3176
              ||||||||    ||||||||:||||:|||||||||||||||||:|||||||||||||||
Db        850 CAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAA 905

Qy       3177 YRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTC 3236
              :::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|||
Db        906 CACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGGGGATTC 965

Qy       3237 CTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYC 3296
              ||||||:|||:|||:||||||||||:|||||||  |||||||:||||:|||||||||||:|
Db        966 CTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAATAGCCAGGTC 1024

Qy       3297 ATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMC 3356
              ||||:|||||:|||||||||||||||||||||||||||||||||||||:|||||||:|:|
Db       1025 ATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACAAC 1084

Qy       3357 TGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTA 3416
              |||||:::|||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1085 TGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGCCCCAGTGTTA 1130

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
              ||:||||||||:|||||||||||||:||:|||::|||||| |||||||||:||:||||||
Db       1131 AGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAGGAATAGCTCT 1189

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
              :|||||||||||||||:|||:|||||:|||||||||||::||||::||||||:||||||
Db       1190 AGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACCTGACTAAGGA 1249

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
              ||:|||||||||||||||||||||:|:|||||||||:|||||||||||||||||||||||
Db       1250 AATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGT 1309

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
              ::|||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       1310 CTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCA 1369

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
              |||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||||
Db       1370 TGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTACT 1429

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
              |||||:|||||||||||||||||:|||||||||||:|||||||||||||||||||||||
```

```
Db       1430 TAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAACTCT 1489

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
Db       1490 TAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACTGTCAACAAGT 1549

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
Db       1550 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGACTGATCCCGA 1609

Qy       3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
Db       1610 CC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGTGGGGT 1668

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db       1669 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 1728

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db       1729 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 1788

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db       1789 CAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 1847

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db       1848 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 1907

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db       1908 GGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACACTGCCGGGGTC 1967

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
Db       1968 ATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATTGAAGCCAAAA 2027

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db       2028 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 2087

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db       2088 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 2147

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
Db       2148 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 2207

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
Db       2208 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 2267

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
Db       2268 GCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 2327

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
Db       2328 AGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTATCGCCAAGCTC 2387

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
Db       2388 CTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTTTACCCA 2447

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
Db       2448 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTTTCACGGGTTG 2507

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
Db       2508 GGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 2567

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
Db       2568 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 2627

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
Db       2628 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 2687

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
Db       2688 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGACA 2747

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
Db       2748 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGCCTATAGCCTT 2807

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
Db       2808 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 2867

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
Db       2868 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 2927

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
Db       2928 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 2987

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
Db       2988 GAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTCCCTTCCCTCT 3045

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
Db       3046 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCA 3105

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
Db       3106 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 3165

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
Db       3166 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 3225

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
Db       3226 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 3266
```

RESULT 9
AAX25656
ID AAX25656 standard; cDNA to mRNA; 2938 BP.
XX
AC AAX25656;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.6A1.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97FR-00008815.

XX
PA (INMR ) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
DR WPI; 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
PS Claim 1; Page 49-52; 106pp; French.
XX
CC This sequence represents clone cl.6A1 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility
XX
SQ Sequence 2938 BP; 878 A; 720 C; 646 G; 692 T; 0 U; 2 Other;

Query Match 36.1%; Score 2706.61; DB 1; Length 2938;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2831; Conservative 90; Mismatches 17; Indels 36; Gaps 15;

```
Qy       586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
             ||:::|||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy       646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
             ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db        61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy       706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy       766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy       826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy       886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
             ||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db       301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy       946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy      1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy      1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
             ||||||||||||:||||::||||||||||||||||||||||||||||||||||||||||||
Db       480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy      1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
             ||||||||||:|||||||||||||||:||||||||||||||||||||||||||||||||||
Db       540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy      1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

Qy      1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
             |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy      1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
             ||||||||||||| | ||  |  |||  |||||||:||||||||||
Db       718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy      1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy      1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
             ||||||||||||||||:||||||||||||||||||:|||||||:||||:||||||||||
Db       833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy      1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952

Qy      1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012

Qy      1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072

Qy      1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132

Qy      1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192

Qy      1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252

Qy      1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312

Qy      1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372

Qy      1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432

Qy      2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492

Qy      2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552

Qy      2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611

Qy      2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
             ||||||| ||||||||||||||||||||||| ||:|||||||||||||||||| |||||
Db       1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668

Qy      2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728

Qy      2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788

Qy       2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
             |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847

Qy       2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906

Qy       2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966

Qy       2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026

Qy       2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086

Qy       2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db       2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||:|||||||
Db       2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             |||||||||||||||||||||||||||||||:||:|::|||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:|||||||||||||||||:||:||||::|||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:|||||||||||:|||||||||||||:||||||||||:|:|:||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:||||||||||||||||    ||||||||:||||:|||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::||||||||||||||||||||||||||||:|||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||||||||||||||||||||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:||||||||:|||||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:|||||||:|||||||||||||||:||||:|||||:||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938

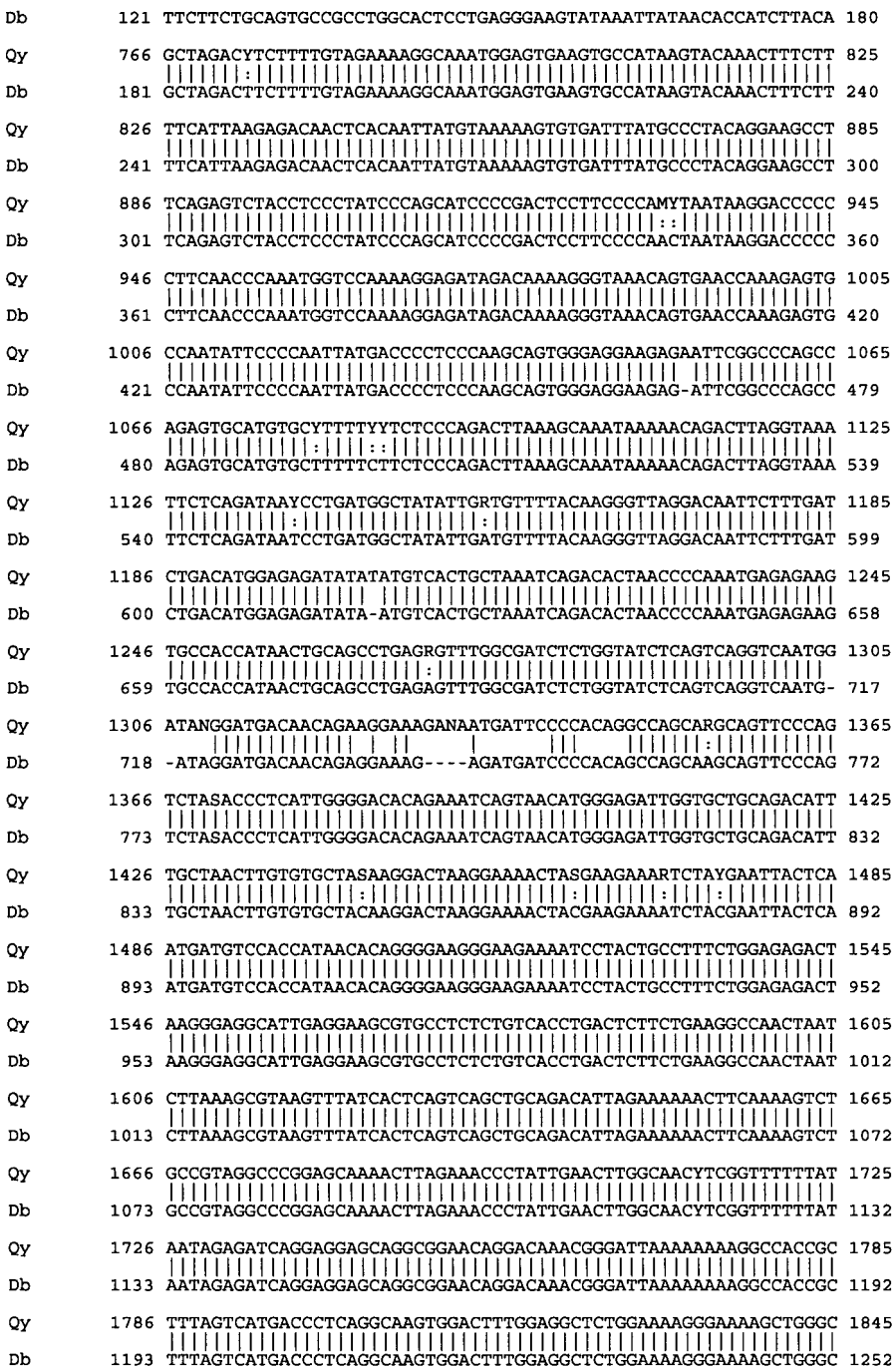

RESULT 10
AAA59206
ID   AAA59206 standard; DNA; 2938 BP.
XX
AC   AAA59206;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Gag and partial pol gene fragment of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 43; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 2938 BP; 878 A; 720 C; 646 G; 692 T; 0 U; 2 Other;

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180
Qy 766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
Db 181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240
Qy 826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
Db 241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300
Qy 886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
Db 301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360
Qy 946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
Db 361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420
Qy 1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
Db 421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479
Qy 1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
Db 480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539
Qy 1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
Db 540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599
Qy 1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
Db 600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658
Qy 1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
Db 659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717
Qy 1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
Db 718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772
Qy 1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
Db 773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832
Qy 1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
Db 833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892
Qy 1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
Db 893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952
Qy 1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
Db 953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012
Qy 1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
Db 1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072
Qy 1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
Db 1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132
Qy 1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
Db 1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192
Qy 1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
Db 1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252
Qy 1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
Db 1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312
Qy 1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
Db 1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372
Qy 1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
Db 1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432
Qy 2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
Db 1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492
Qy 2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
Db 1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552
Qy 2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
Db 1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611
Qy 2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
Db 1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668
Qy 2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
Db 1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728
Qy 2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
Db 1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788
Qy 2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
Db 1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847
Qy 2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
Db 1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906
Qy 2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
Db 1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966
Qy 2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
Db 1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026
Qy 2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
Db 2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086
Qy 2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
Db 2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146
Qy 2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
Db 2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206
Qy 2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
Db 2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266
Qy 2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
Db 2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

```
Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             ||||||||||||||||||||||||||||||:||:|::|||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:|||||||||||:|||||||||||:|||||||||:|:|:||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:||||||||||||||||    ||||||||:||||:||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::||||||||||||||||||||||||||:||||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||  |||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC-------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:||||||||:|||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:||||||:|||||||||||||||:|||:|||||:||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938
```

```
RESULT 11
AAX25663
ID   AAX25663 standard; cDNA to mRNA; 3372 BP.
XX
AC   AAX25663;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.Pi5T.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 64-67; 106pp; French.
XX
CC   This sequence represents clone cl.Pi5T of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 U; 0 Other;

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;
```

```
Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||:|||||||||||||||||||||||||:||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             ||||||||||||||||||||||||||||||:||:|::|||||||||||||  ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:||||||||||:|||||||||||||:||||||||||:|:|:||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||    ||||||||:||||:|||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::||||||||||||||||||||||||||||||:|||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

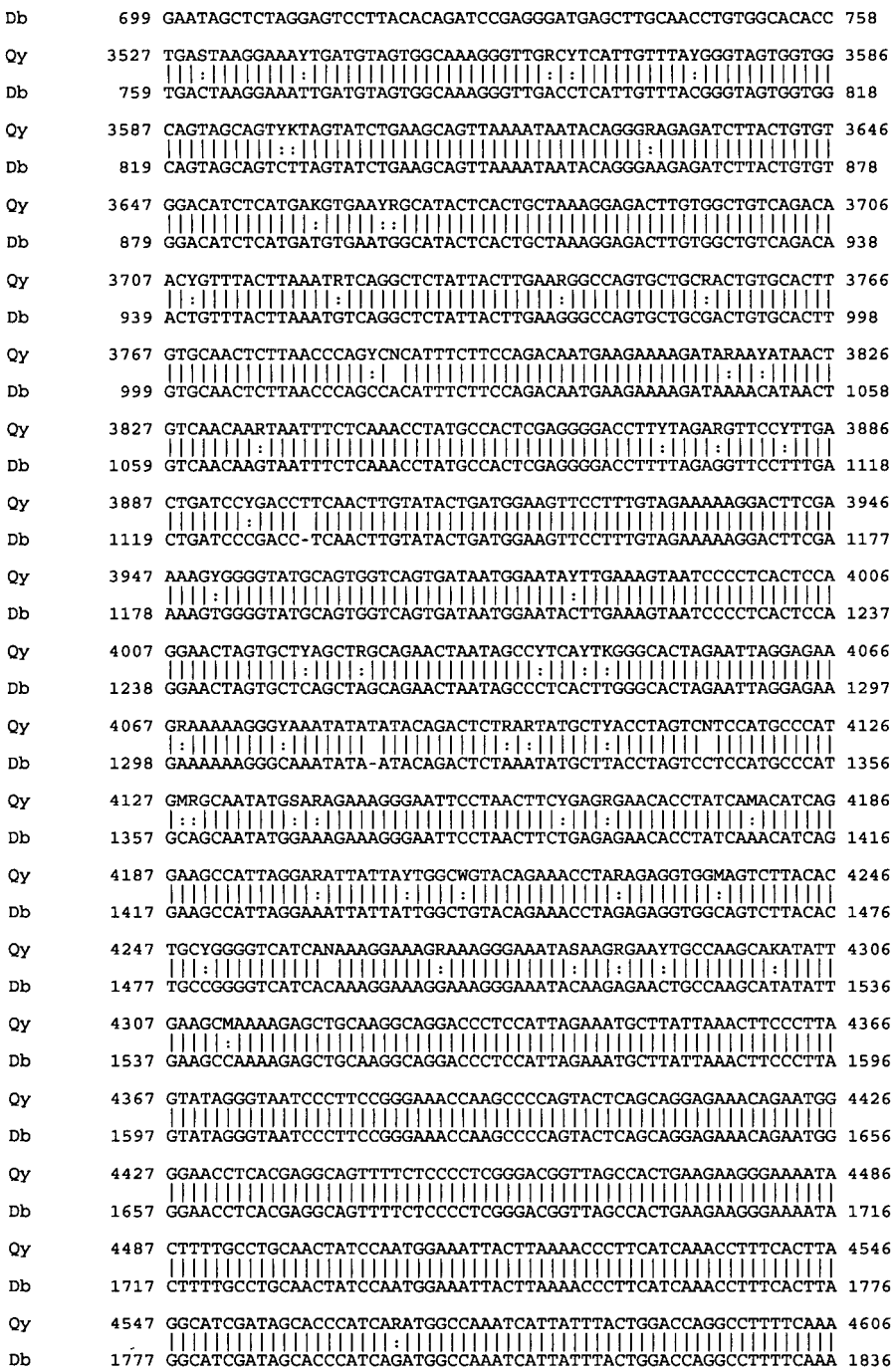
Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::|||||||||:|||:|||:||||||||||:|||||| |||||||:||||:|||
Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db        534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
             ||||||:|:||||||:::|||||||||||||||||||  |||||||:
Db        594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC-------------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
             ||||||||||||:||||||||:||||||||||||||:||:|||::|||||| |||||||||
Db        640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
             :||:|||||:|||||||||||||||:|||:|||||:||||||||||::||||::|||
```




RESULT 12
AAA59213
ID AAA59213 standard; DNA; 3372 BP.
XX
AC AAA59213;
XX
DT 07-NOV-2000 (first entry)

```
XX
DE   Partial pol gene and U3-R region sequences of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 47-48; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 U; 0 Other;

  Query Match           35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:|||||||||||||||||||||||:||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||:||:|::|||||||||||| ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:|||||||||||:|||||||||||||||||:|:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||||    ||||||||:||||:||||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||:||||||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::|||||||||:|||:|||:||||||||||:||||||| |||||||:||||:|||
Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db        534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
             ||||||:|:||||||:::|||||||||||||||||||                ||||||||:
Db        594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
             ||||||||||||:||||||||:|||||||||||||:||:|||::|||||| ||||||||
Db        640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
             :||:||||||:||||||||||||||||:|||:|||||:||||||||||::||||::|||
Db        699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
             |||:||||||||:|||||||||||||||||||||:|:||||||||||:||||||||||||
Db        759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
             ||||||||||::|||||||||||||||||||||||||||||:|||||||||||||||||
Db        819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
             |||||||||||||:|||||::|||||||||||||||||||||||||||||||||||||||
Db        879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
             ||:|||||||||||:|||||||||||||||||||:||||||||||:||||||||||||||
Db        939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
             |||||||||||||||||:| ||||||||||||||||||||||||||:||:|||||||
Db        999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||||:|||||||||||||||||||||||||||||||||||||:||||:|||||||||
Db       1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
             ||||:||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
             ||||||||||||:|||:|||||||||||||||||||:|||:|:||||||||||||||||||
Db       1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
             |:||||||||:||||||| ||||||||||:|:||||||:||||||| ||||||||||||
Db       1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356

Qy       4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
             |::|||||||:|:|||||||||||||||||||||||:|||:||||||||||:|||||||
Db       1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416

Qy       4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
             ||||||||||||||:||||||:||||:|||||||||||||:||||||||:||||||||||
Db       1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476
```

```
Qy       4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
             |||:|||||||||| |||||||||:||||||||||:|||:|||:|||||||||:|||||
Db       1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536

Qy       4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596

Qy       4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656

Qy       4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716

Qy       4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:|||:|||||||||||||:|||||||:|:|||||||||||||||||:||
Db       1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
             ||||||||||||||||||:|||||||:|||||||||||||||||:|:||:||||||||||
Db       1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956

Qy       4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016

Qy       4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076

Qy       4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136

Qy       4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196

Qy       4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256

Qy       5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db       2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316

Qy       5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376

Qy       5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496

Qy       5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
             |||||||||||||||  |||  | ||||||||::|||||||||||||||||||||||||
Db       2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554

Qy       5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785
```

RESULT 13
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
AC ABN97927;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HERV 7 env.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
OS Human endogenous retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Claim 1; Page 128-129; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
SQ Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 U; 0 Other;

Query Match 34.1%; Score 2551.4; DB 1; Length 2599;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2552; Conservative 39; Mismatches 5; Indels 2; Gaps 2;

```
Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714
             |||   | | ||||||||||||||||||:||||||:||||||||||||||||:|:||
```

```
Db        4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy     4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123

Qy     4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
            |||||:||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db      124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183

Qy     4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243

Qy     4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303

Qy     4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363

Qy     5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
            |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423

Qy     5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
            |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483

Qy     5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543

Qy     5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 603

Qy     5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
            ||||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db      604 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 663

Qy     5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
            |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db      664 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 723

Qy     5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
            ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db      724 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 783

Qy     5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      784 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 843

Qy     5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      844 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 903

Qy     5555 AAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTG 5614
            |||||||||||||||||||| ||||||||:|||||||||||||||||||:||:||
Db      904 AAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTG 962

Qy     5615 TTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 5674
            ||:|||:||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db      963 TTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 1022

Qy     5675 CCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 5734
            |||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1023 CCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 1082

Qy     5735 ATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 5794
            ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db     1083 ATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 1142

Qy     5795 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 5854
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1143 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 1202

Qy     5855 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 5914
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1203 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 1262

Qy     5915 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 5974
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1263 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 1322

Qy     5975 TAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 6034
            ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db     1323 TAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 1382

Qy     6035 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 6094
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1383 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 1442

Qy     6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 6154
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 1502

Qy     6155 TGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 6214
            |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1503 TGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 1562

Qy     6215 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATA 6274
            ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db     1563 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATA 1622

Qy     6275 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 6334
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1623 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 1682

Qy     6335 TCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 6394
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1683 TCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 1742

Qy     6395 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 6454
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1743 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 1802

Qy     6455 TCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 6514
            ||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db     1803 TCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 1862

Qy     6515 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 6574
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1863 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 1922

Qy     6575 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 6634
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1923 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 1982

Qy     6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1983 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 2042

Qy     6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754
            ||||||||||||||||||||||||||:||||||||||||||||||:|||||||||||||||
Db     2043 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 2102

Qy     6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814
            |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2103 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 2162

Qy     6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874
            ||||||||||||||||||:|||||||||:|||||||||||||||||:|||||:||||
Db     2163 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 2222
```

```
Qy       6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       2223 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 2282

Qy       6935 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCC 6994
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       2283 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCC 2342

Qy       6995 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGC 7054
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       2343 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGC 2402

Qy       7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATG 7114
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATG 2462

Qy       7115 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 7174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2463 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 2522

Qy       7175 CCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGT 7234
             |||||||||||||||||||||||||||: ||||||||||||||||||||||||||||||
Db       2523 CCAATTCAGCAGGAAGCAGTTAGAGCGGTC-TCGGCCAACCTCCCCAACAGCACTTAGGT 2581

Qy       7235 TTTCCTGTTGAGATGGGG 7252
             ||||||||||||||||||
Db       2582 TTTCCTGTTGAGATGGGG 2599
```

RESULT 14
ABN97930
ID ABN97930 standard; DNA; 2784 BP.
XX
AC ABN97930;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HE2.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
OS Human endogenous retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Claim 3; Page 133; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
SQ Sequence 2784 BP; 816 A; 755 C; 535 G; 678 T; 0 U; 0 Other;

Query Match 30.9%; Score 2315.4; DB 1; Length 2784;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 2523; Conservative 39; Mismatches 180; Indels 66; Gaps 9;

```
Qy       4674 CTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTA 4732
             ||||||||||:||||||:||||||||| |||||  :|:||:|||||||||| |||||||
Db          1 CTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAAGAGAAGACTGGCAACTAGATTTTA 60

Qy       4733 CCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGG 4792
             |||| | ||||||| |||||||||||||||||||||||| ||||||:|||||||||| |
Db         61 CCCATATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGATACTTTCACTG 120

Qy       4793 GTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAG 4852
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||| |
Db        121 GTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAA------CG 174

Qy       4853 TTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTG 4912
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||
Db        175 TTCATGAAATAATTCCCAGATTCGGACTTCCCCAAGGCTTACAGAGTGACAATGGCCCTG 234

Qy       4913 CTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTAC 4972
             ||||| |||| |||||||||| |||||||||||| |||||||||| |||||||| ||
Db        235 CTTTCAAGGCTACAGTAACCCAAGGAGTATCCCAGGTGTTAGGTATACAATATCACTCAC 294

Qy       4973 ACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAG 5032
             |||||||||| |||||||||||||||||| ||||| |||||||||| |||:|||||||
Db        295 ACTGCGCCTGGAGGCCACAGTCCTCAGGAAAGGTGGAGAAAATGAACAAAACACTCAAAT 354

Qy       5033 GACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAG 5092
             |||||||||||||| || |||||||||||||||| |||||||||:||||||||||||||
Db        355 GACATCTAAAAAAGCTAATCCAGGAAACCCACCTCGCATGGCCTGCTCTGTTGCCTATAG 414

Qy       5093 CCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152
             |||||  ||||||| | |||| |||||||||||||||||||| |||||| |||||||||
Db        415 CCTTACTAAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGTCCATACAAAATGCTGT 474

Qy       5153 ATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTG 5211
             ||||| |||||||| |||||||||| |||| |||||||| | ||||||||||||||||||
Db        475 ATGGACGGCCCTTCCTAACCAATGAACTTGGGCTTGACCGAGAGACAGCCAACTTAGTTG 534

Qy       5212 CAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATC 5271
             |||||||| |||||||||||||||||||| |||||||||||||||||| ||| ||| ||
Db        535 CAGACATCATCTCCTTAGCCAAATATCAACAGGTTCTTAAAACATTACAGGGAGCCTGTC 594

Qy       5272 CCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTC 5331
             ||  |||||||||||| |||||||||||||: |||||||||||||||||||||||||||||
Db        595 CCCAAGAAGAGGGAAAGGAACTATTCCACCCTGGTGACATGGTATTAGTCAAGTCCCTTC 654

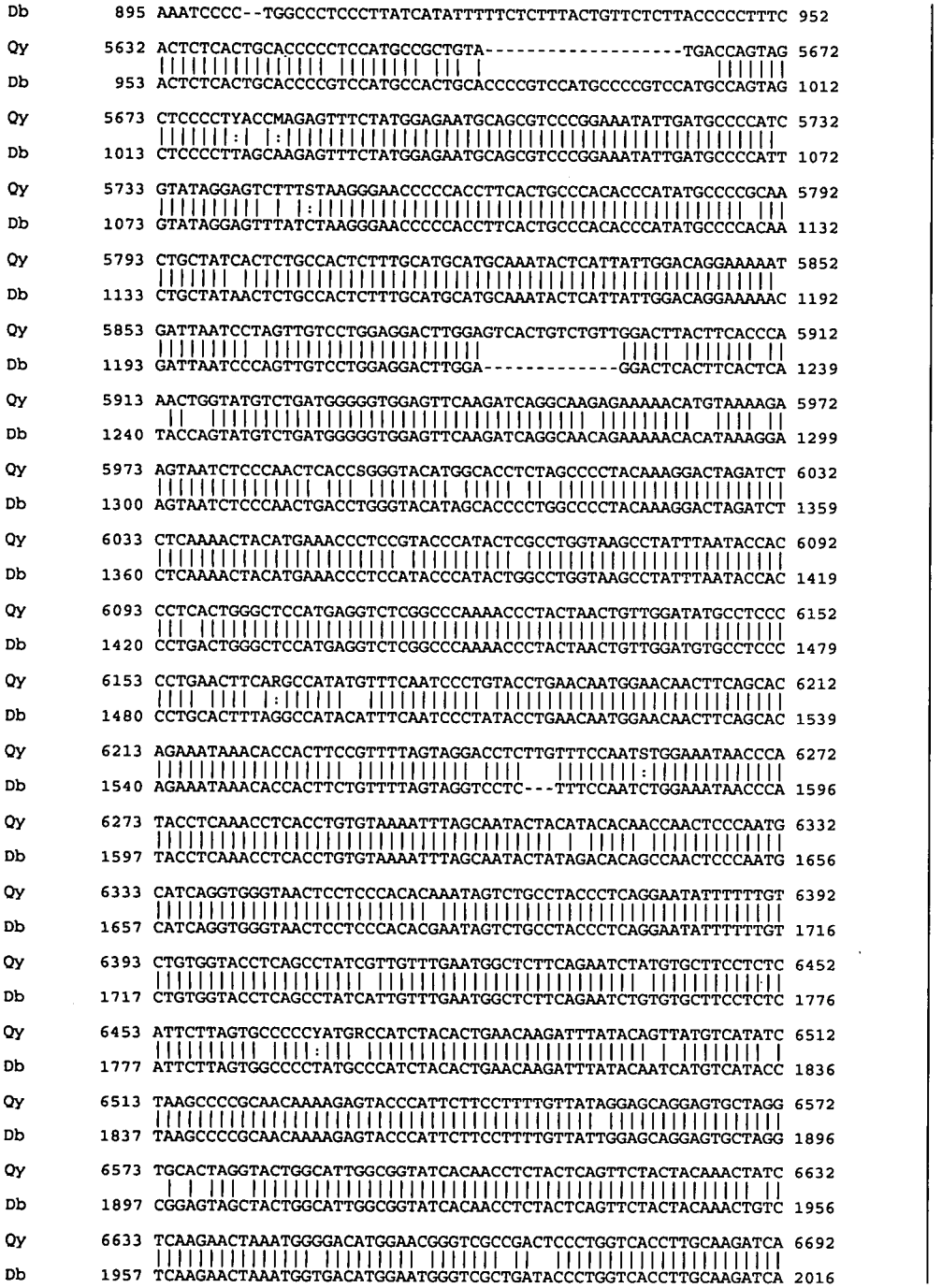
Qy       5332 YCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTA 5391
             :|||||||||||||||||||||||||||||||||||||||  |||||||| |||||||:||
Db        655 CCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGAAACTACCCAGCCATTTTATCTA 714

Qy       5392 CCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAT 5451
             ||| ||| || |||||||||||||||| ||||||||||||||||||||  |||||||
Db        715 CCCTAACGGCAGTTAAAGTGGCTGGAGCGGAGTCTTGGATACATCACACTCAAGTCAAAC 774

Qy       5452 CCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAAC 5511
             ||||||||||||||||||||  |||||||| ||||||| ||||||||||||||||||||
Db        775 CCTGGATACTGCCAAAGGAACTCAAAAATCCATGAGACAATGCTAGCTATTCCTGTGAAC 834

Qy       5512 CTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCAT 5571
             |||||||||| |||||||||||||||  |||||||| |||||||||||||||||||| |
Db        835 CTCTAGAGGATCTGCGCCTGCTCTTCAAATGACAACCAGGGGGAAAGTAACTAAAATCGT 894

Qy       5572 AAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTC 5631
             ||||||||  |||:||||||||||||||||||||:||:||||:| |:|||| :||||
```




RESULT 15
AAX25664
ID AAX25664 standard; cDNA to mRNA; 2372 BP.
XX
AC AAX25664;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.44.4.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.

```
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 68-70; 106pp; French.
XX
CC   This sequence represents clone cl.44.4 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 2372 BP; 674 A; 593 C; 544 G; 557 T; 0 U; 4 Other;

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
               ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||||||::||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              ||||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              |||||||   ||| | |     ||||| ||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:||||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
             ||||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

RESULT 16
AAA59214
ID AAA59214 standard; DNA; 2372 BP.
XX
AC AAA59214;
XX
DT 07-NOV-2000 (first entry)
XX
DE R-U5 region and partial pol gene sequences of HERV-W from human genome.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200043521-A2.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-FR000144.
XX
PR 21-JAN-1999; 99FR-00000888.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Mallet F, Voisset C;
XX
DR WPI; 2000-499229/44.
XX
PT New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy, contains
PT at least part of the gag gene.
XX
PS Disclosure; Page 49; 53pp; French.
XX
CC The present sequence represents an endogenetic retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived from
CC a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC retrovirus is associated with autoimmune disease, failure of pregnancy or
CC disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC from it, are useful for diagnosis of autoimmune disease (specifically
CC multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC fragments may also be used for in situ labelling of isolated chromosomes,
CC while the transcription product can be used to study or monitor T cell
CC proliferation in vitro
XX
SQ Sequence 2372 BP; 674 A; 593 C; 544 G; 557 T; 0 U; 4 Other;

Query Match 30.1%; Score 2256.8; DB 1; Length 2372;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2332; Conservative 30; Mismatches 10; Indels 10; Gaps 8;

```
Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              ||||||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              ||||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||||||::||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:|||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              |||||||||||||:||||||||  ||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              |||||||||||||||||||||||||||||||||||||:||||||||||||||:||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              ||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:||||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
              |||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

```
RESULT 17
AAX25661
ID   AAX25661 standard; cDNA to mRNA; 2782 BP.
XX
AC   AAX25661;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.PH74.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
```

```
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 60-63; 106pp; French.
XX
CC   This sequence represents clone cl.PH74 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||:||||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764
```

```
RESULT 18
AAA59211
ID   AAA59211 standard; DNA; 2782 BP.
XX
AC   AAA59211;
XX
DT   07-NOV-2000  (first entry)
XX
DE   5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 46-47; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

  Query Match            29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity  98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764
```

```
RESULT 19
AAH20069
ID   AAH20069 standard; DNA; 2782 BP.
XX
AC   AAH20069;
XX
DT   08-AUG-2001  (first entry)
XX
DE   HERV-W envelope protein G encoding nucleic acid.
XX
KW   Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW   envelope protein; multiple sclerosis-related superantigen; vaccine;
KW   surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW   antisense-therapy; autoimmune disorder; ds.
XX
OS   Human endogenous retrovirus.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .762
FT                   /*tag=  a
FT   CDS             763. .2379
FT                   /*tag=  b
FT                   /product= "HERV-W envelope protein G"
FT                   /transl_except= (pos:790. .792,aa:Phe)
FT                   /transl_except= (pos:793. .795,aa:Thr)
FT                   /transl_except= (pos:812. .814,aa:Leu)
FT                   /transl_except= (pos:818. .820,aa:Ser)
FT                   /transl_except= (pos:862. .864,aa:Tyr)
FT                   /transl_except= (pos:865. .867,aa:Gln)
FT                   /transl_except= (pos:1174. .1176,aa:Arg)
FT                   /transl_except= (pos:1441. .1443,aa:Leu)
FT                   /transl_except= (pos:1903. .1905,aa:Gln)
FT                   /transl_except= (pos:2017. .2019,aa:Lys)
FT                   /transl_except= (pos:2026. .2028,aa:Arg)
FT                   /transl_except= (pos:2044. .2046,aa:Arg)
FT                   /transl_except= (pos:2089. .2091,aa:Gln)
FT                   /transl_except= (pos:2170. .2172,aa:Asn)
FT   3'UTR           2380. .2782
FT                   /*tag=  c
XX
PN   WO200131021-A1.
XX
PD   03-MAY-2001.
XX
PF   30-OCT-2000; 2000WO-EP010659.
XX
PR   28-OCT-1999;   99EP-00402690.
XX
PA   (UYGE-) UNIV GENEVE.
XX
PI   Conrad B,  Mach B;
XX
DR   WPI; 2001-316336/33.
DR   P-PSDB; AAB75138.
XX
PT   New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT   activity useful for diagnosing and treating multiple sclerosis.
XX
PS   Claim 13; Fig 9; 94pp; English.
```

XX
CC On the basis of the PBS t-RNA motif used for the classification of human
CC endogenous retrovirus (HERVs) the full length endogenous provirus which
CC was been located on the long arm of human chromosome 7 (7q21-22) has been
CC designated HERV-W. The present invention describes proteins or peptides
CC (I) having superantigen (SAg) activity comprising the ENV protein (ENV)
CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC have neuroprotective activity, and can be used in: vaccines; antisense-
CC therapy; and HERV-W SAg activity-inhibitors. (I) and encoding DNA/RNA are
CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC disorders. (I) are also useful for identifying substances (and optionally
CC recovering) capable of binding to a retroviral superantigen associated
CC with MS, substances capable of blocking SAg activity and substances
CC capable of blocking transcription or translation of HERV-W retroviral
CC superantigen. A protein or peptide derived from (I), modified to be
CC devoid of SAg activity and being capable of generating an immune response
CC against HERV-W retroviral SAg is useful in therapy. Nucleic acid
CC molecules encoding (I) are useful as vaccines against MS. Substances
CC capable of blocking SAg activity, capable of binding to a retroviral
CC superantigen associated with MS, or capable of blocking transcription or
CC translation of HERV-W retroviral superantigen for use in treating or
CC preventing MS, obtained using (I) are useful for the treatment and
CC prevention of MS. (I) and nucleic acids encoding them are useful for
CC diagnosing autoimmune disease. The present sequence encodes the
CC specifically claimed envelope protein of HERV-W designated G
XX
SQ Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;

Query Match 29.4%; Score 2205.1; DB 1; Length 2782;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 2194; Conservative 36; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
```

```
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
              ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
              ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
              ||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
              |||||:||||:
Db       2754 TTGCAACTGCA 2764


RESULT 20
AAF55630
ID   AAF55630 standard; DNA; 2781 BP.
XX
AC   AAF55630;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of a human endogenous retrovirus envelope protein.
XX
KW   Envelope protein; HERV; syncytia formation; placental development;
KW   syncytia; cancer; cell adhesion; ss.
XX
OS   Human endogenous retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             762. .2378
FT                   /*tag= a
FT                   /product= "envelope protein"
XX
PN   WO200116171-A1.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-FR002429.
XX
PR   01-SEP-1999;   99FR-00011141.
PR   15-SEP-1999;   99FR-00011793.
XX
PA   (INMR ) BIO MERIEUX.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Mallet F,  Cosset F,  Blond J,  Lavillette D,  Bouton O,  Ruggieri A;
XX
DR   WPI; 2001-226676/23.
DR   P-PSDB; AAB67652.
XX
PT   Detecting expression of human endogenous retrovirus envelope protein in
PT   cells of a tissue or culture, from its ability to induce syncytia.
XX
PS   Disclosure; Page 44-45; 57pp; French.
XX
CC   The present sequence encodes a human endogenous retrovirus envelope
CC   protein. The specification describes a method for detecting expression of
CC   an envelope protein from a human endogenous retrovirus (HERV), in cells,
CC   of a tissue or culture. The method comprises detecting syncytia formation
CC   due to the fusogenic properties of the envelope protein. Envelope
CC   polypeptides and polynucleotides are used to produce therapeutic or
CC   prophylactic compositions, particularly for treatment of cancer, to
CC   correct defects in placental development (or other natural formation of
CC   other types of syncytia), and to promote adhesion of cells in grafts or
CC   cellular repair processes. Expression of sequences antisense to the
CC   polynucleotide are used to prevent formation of syncytia
XX
SQ   Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

  Query Match             29.4%;  Score 2201.9;  DB 1;  Length 2781;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2192;  Conservative   36;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        534 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 593

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 653

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 713

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
              |||||||||||||||||||||||||||||||||||||||||| |||||||:||||||
Db        714 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 772

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
              ||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||
Db        773 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 832

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
              ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db        833 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 892

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        893 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 952

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1012

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1072

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1132

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1133 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1192

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1252
```

```
Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1312

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1313 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1372

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1432

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1492

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1552

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1612

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db       1613 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1672

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1732

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 TATAGGAGCGGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1792

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1852

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             |||||||||||||||||||| |||||||||||||||||||||||||||:|||||||||
Db       1853 CCTGGTCACCTTGCAAGATCAGCTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1912

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1913 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1972

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1973 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2032

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||
Db       2033 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2092

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2152

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||||
Db       2153 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2212

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2272

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2332

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2333 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2392

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2452

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:||||||||:||||||||:|||||||||||||||||||||||
Db       2453 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2512

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2513 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2572

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||||
Db       2573 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2632

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2633 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2692

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2752

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2753 TTGCAACTGCA 2763
```

RESULT 21
AAX25668
ID  AAX25668 standard; cDNA to mRNA; 2575 BP.
XX
AC  AAX25668;
XX
DT  21-MAY-1999  (first entry)
XX
DE  Human endogenous retrovirus W clone cl.7A21.
XX
KW  Clone; human endogenous retrovirus; genome; autoimmune disease;
KW  multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW  disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS  Human endogenous retrovirus.
XX
PN  WO9902696-A1.
XX
PD  21-JAN-1999.
XX
PF  06-JUL-1998;  98WO-FR001442.
XX
PR  07-JUL-1997;  97FR-00008815.
XX
PA  (INMR ) BIO MERIEUX.
XX
PI  Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR  WPI; 1999-120897/10.
XX
PT  New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT  exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT  disease, and abnormal or failed pregnancy.
XX
PS  Claim 1; Page 80-83; 106pp; French.
XX
CC  This sequence represents clone cl.7A21 of the human endogenous retrovirus
CC  (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC  by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC  rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-

```
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 2575 BP; 817 A; 608 C; 546 G; 602 T; 0 U; 2 Other;

  Query Match             28.6%;  Score 2141.4;  DB 1;  Length 2575;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches 2275;  Conservative  134;  Mismatches  159;  Indels   20;  Gaps   11;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTG-AGYCARTYMTCATAC 2969
Db          1 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAAGCCAATTCTCATAC 60

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
Db         61 CTGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACC 120

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
Db        121 TTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTT 180

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db        181 TCCAAACCAAAGACCCAGCTCTGCTCACAGCAGGTTAAATACTTGGGGCTAAAATTATCC 240

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db        241 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCAAGCCTATACTGGCTTATCCTCATCCC 300

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        301 CAAATCCTAAAGCAACTAAGAGAGTTCCTTAGCATAACAGGTTTCTGCTGAATATGGATT 360

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        361 -CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGC 419

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        420 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAG 477

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        478 AGGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTAT 537

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        538 GTCACAGAAAAAACAGGAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTT 597

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        598 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 657

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        658 TTTATGGGTAGTGGCGGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATACAGGG 717

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        718 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 777

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
Db        778 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGT 837

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
Db        838 GCAGCAACTGCGCAGTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATG---- 893

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869

Db        894 -AAGATAGAACATAACTGCCAACAAGTAATTTCTCAAACCTAGGCCGCTCGAGGGAACCT 952

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
Db        953 TTTAGAGGTTCCCTTAACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 1011

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
Db       1012 TAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGCTCAGTGATAATGGAATACTTGAAA 1071

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
Db       1072 ATAATCCCTTCATTCCAGGAACCAGCGTTCAGCTGGCAGAATTAATAGCCCTCACTCGGG 1131

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
Db       1132 CATTAGAATTAGGAGAAGGAAAAAGGGTAAATACACATACAGATTCTAAGTATGTTTACT 1191

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
Db       1192 TAGTCCTCCGTGCCCACGCAGCAATATGGAGAGAAAGGGAATGCTTAACTTCTGAGGGAA 1251

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229
Db       1252 CACCTATCAAACATCAGGAAGTTATTAGGAGATTATTATTGGCTATACAGAAACCTAAAG 1311

Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289
Db       1312 AGGTGGCAGTCTTACACTGCTGGGGTGGTCAGAAAGAAAAGGAAAGGGAAATAAAAGGGA 1371

Qy       4290 AYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGC 4349
Db       1372 ACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGC 1431

Qy       4350 TTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAG 4409
Db       1432 TTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAATACTTAG 1491

Qy       4410 CAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAG 4466
Db       1492 AAAAAGAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCTCCCCTCAGGATGGCTAG 1551

Qy       4467 CCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCC 4526
Db       1552 CCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCC 1611

Qy       4527 TTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTA 4586
Db       1612 TTCACCAAACCTTTCACTTAGACATTGATAGCACCCATCAGATGGCCAAATCATTATTTA 1671

Qy       4587 CTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARA 4646
Db       1672 CTGGACCAGGCCTTTTCAAAACTATCAAGCAGCTAGTCAGGGCCTGTGAAGTGTGCCGAA 1731

Qy       4647 RAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACC 4706
Db       1732 GAAATAATCCCATGCCTTATCACCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACC 1791

Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
Db       1792 CAGGAGAAGRVTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 1851

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
Db       1852 TACTAGTTTGGGTAGATACTTTCACTGGTTGGGCAGAGACCTTCCCCTGTAAGACAGAAA 1911

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
Db       1912 AGTCCCAAGAGGTAATAAAGGCATTAGTTCATGAAATAATTCCCAGATTCAGACTTCCCT 1971

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
Db       1972 GAGGCTTACAGAGTGACAATGGCCCTGCTTTCAAGGCTACAGTAACCCA-GGAGTATCCC 2030
```

```
Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
             ||| |||||||||| |||||||||||||||||||| |||   ||||||||||||||||
Db       2031 AGGTGTTAGGTATACAATATCACTTACACTGCGCCTGGAGG---CAGTCCTCAGGGAAGG 2087

Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCT-AAAAAAGCAAACCCAGGAAACCCAC 5064
              ||||||| ||||||||:|||||||| ||||||| |||||||| |||||||||| |||| 
Db       2088 CCGAGAAACTGAATGAAACACTCAAACGACATCTAAAAAAAGCTAACCCAGGAAAACCAC 2147

Qy       5065 CTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAA 5124
              ||||||||||||:|||||||||||||||||||| |||||||   |||| ||||||||||
Db       2148 CTCACATGGCCTGCTCTGTTGCCTATAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAA 2207

Qy       5125 GCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 5184
              |||||||||||||||||||||||||| |||||| ||||||| ||||||||||||||||
Db       2208 GCAGGACTTAGCCCATACGAAATGCTATATGGATAGCCCTTCCTAACCAATGACCTTGTG 2267

Qy       5185 CTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAA 5243
              ||||||  | ||| ||||||||||||||||||||||||||||||| ||||||||||||||
Db       2268 CTTGACTGAGAGAGAGCCAACTTAGTTGCAGACATCACCTCCTTATCCAAATATCAACAA 2327

Qy       5244 GTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCW 5303
              |||||||||||||||||||| ||| |||| |||||||||| || ||||||||||||||:
Db       2328 GTTCTTAAAACATTACAAGGAGCCTGTCCCCGAGAAGAGGGGAAGGAACTATTCCACCCT 2387

Qy       5304 WGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGG 5363
              ||||||||||||||||||||||||||:||||||||| |  | ||||||| |||||||
Db       2388 GGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCTCATTGCCTAGATATATCCTGG 2447

Qy       5364 GAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAG 5423
              |||||||||||||||||||||||:||||||||| || || ||||||||||||||||||
Db       2448 GAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGCAGTAAAAGTGGCTGGAGTGGAG 2507

Qy       5424 TCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCA 5483
              |||||||||||||||||| |||||||| |||||||| | ||||||||||||||||||||
Db       2508 TCTTGGATACATCACACTCGAGTCAAACCCTGGATATTACCAAAGGAACCTGAAAATCCA 2567

Qy       5484 GGAGACAA 5491
              ||||||||
Db       2568 GGAGACAA 2575
```

```
RESULT 22
AAX25667
ID   AAX25667 standard; cDNA to mRNA; 2585 BP.
XX
AC   AAX25667;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.7A20.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 77-80; 106pp; French.
XX
CC   This sequence represents clone cl.7A20 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 2585 BP; 821 A; 623 C; 542 G; 597 T; 0 U; 2 Other;

  Query Match             27.2%;  Score 2036.1;  DB 1;  Length 2585;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches 2233;  Conservative  129;  Mismatches  213;  Indels   19;  Gaps   14;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
              |||||||||:|||||||||||||||||| ||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
              || |||||||||||||||||| :|||:|||||||| |||||||:||: ||::||||||||
Db         61 ACTTGGACACTCTTGTCCTTTGGTATGTGGATGATCTACTTTTAGCCACCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
              |||||||||||||||||| ||||| :|||||:|| ||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCAACCAAGTGCTCTTAAACTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
              |||||||||:|:|:|||||:||||||| |||||||||||||||||||||:||||:|||||
Db        181 TTTCCAAACCAGAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
              ||||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
              :||||||||| |||||| ||||:|::|||||||||:||| |||:| |||  |||:|||||
Db        301 CCAAAACCCTGAAGCAATTAAGAGGGTTCCTTGGCATAAAAGGCTGCTGTTGAATATGGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
              || |||||||:    :||||||||||:|||||:| |||:|||||| || ||||||||||
Db        361 TT-CCCAGGTACAATGAAATAGCCAGGCCATTATACACACTAATTACGGGAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
              |||||||||||||||:||  |||||:|:||||||:::||| |||||||||||||||  ||
Db        420 GCCAATACCCATTTAGTAGAATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAG 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  || |||||:|||||||||||||:|||||||  :| ||||||||||||:||:||
Db        480 AAGGCCCT--AATCCAAGCCCCAGTGTTAAGCTTGCCAATGGAGCAAGACTTTTCTTTAT 537

Qy       3448 AYRTCACAG--AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
              |::||||||  ||||||||||:||:||||||:| |||||||||||||:|||:|||||: |
Db        538 ATGTCACAGAAAAAAAAACAGGAATAGCTCTAGAAGTCCTTACACAGGTCCGAGGGACCA 597

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
              |||| |||| ::||||::||||||:||||||||:|||||||||||||||||||||: :||
Db        598 GCTTACAACACATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGACTC 657

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
              ||||||||: |||||||| ||||||||||::||| |||||||||||||||| ||||
Db        658 ATTGTTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATAC 717

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
              ||||:||| |||||||||||||||||||||||:|||||::|||||||||||||||||||
Db        718 AGGGAAGANATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAG 777

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
```

```
Db        778 GAGAC-TGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATCACTTGAANGGC 836

Qy       3746 CAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATG 3805
Db        837 CAGTGCTGCCACTGTGCACTTGTGCAACTCTTAACCCACCCACATTTCTTCCAGACAATG 896

Qy       3806 AAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
Db        897 AAGAAAAGATAGAACATAACTGTCAACAAGTGATTGTTCAAACCTACACCGCTCGAAGGG 956

Qy       3866 ACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCC 3925
Db        957 ACCTTCTAGAGGTTCCCTTGACTGATCCTGAGC-TCAACTTCTATACTGATGGAAGTTCC 1015

Qy       3926 TTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTT 3985
Db        1016 TTTTGTAGAAAAAGGACTTCGAAAGGCGGGTATGCAGTGGCCAGTGATAATGGAATACTT 1075

Qy       3986 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYT 4045
Db        1076 GAAAGTAATCCCTTCACTCCAGAAACTAGCATTCAGCTGGCAGAATTAATAGCCTTCACT 1135

Qy       4046 KGGGCACTAGAA-TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGC 4104
Db        1136 TGGGCATTAGAACACAGGAGAAGGAAAAGGAGTAAATATATATACAGACTCCAAGTATGC 1195

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
Db        1196 TTACTTAGTCCTCCATGCCCATGCAGCAATATAGAGAGAAAGCGAATTCCTAACTTCTGA 1255

Qy       4165 GRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACC 4224
Db        1256 GGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTACTGGCTGTACAGAAACC 1315

Qy       4225 TARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASA 4284
Db        1316 T--AGAGGTGGCAGTCTTACATGGCCGAGATCATCAGAAAGGAAAAGAAAGGGAAATAGA 1373

Qy       4285 AGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGA 4344
Db        1374 AGGGAACTGCCAAGTGGATATTGAAGCCAAAAGAGCTGCAAGGCGGGACCCTCCATTAGA 1433

Qy       4345 AATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTA 4404
Db        1434 AATGCTTATAGAAGGACCCCTAGTACAGGGCAATCCCCTTCAGGAAACCAAGCCCCAATA 1493

Qy       4405 CTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACG 4461
Db        1494 CTCAGCAGAAGAAATGGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATG 1553

Qy       4462 GTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAA 4521
Db        1554 GCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAA 1613

Qy       4522 AACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATT 4581
Db        1614 AACCCTTCACCAAACCTTTCGCTTAGGCATTGATAGCACCCATCAGATGGCTAAATCATT 1673

Qy       4582 ATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTG 4641
Db        1674 ATTTACTAGACCACACCTTTTCAAAACTATCAAGCAGACAGTTAGGGCCTGTGAAGTGTG 1733

Qy       4642 CCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCA 4701
Db        1734 CCAAAGAAATAATCCCCTGCCTTATCGCCAAACTCCTTCAGGAGAAAAAAGAACAGGCCA 1793

Qy       4702 TTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCA 4760
Db        1794 TTACCCAGGAGAAGAGTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCA 1853

Qy       4761 GTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGAC 4820
Db        1854 GTATCTACTAGTCTGGGTAGATACTTTCACTGGTTGGGCGGAGGCCTTCCCTTGTAGGAC 1913

Qy       4821 AGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACT 4880
Db        1914 AGAACAGGCCCATGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTTGGATT 1973

Qy       4881 TCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGT 4940
Db        1974 TCCCCAAGGCTTACAGAGTGATAACGGCCCCACTTTCAAGGCTACAGTAACCCAGGGAGT 2033

Qy       4941 ATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGG 5000
Db        2034 ATCCCAGACATTAGACATACAATATCACTTACACTGAGCCCGGAGGCCACAATCCTCA-G 2092

Qy       5001 GAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAAC 5060
Db        2093 GAAAGTTGAGAAAATGAATGAAACGCTCAAATGACATCTAAAAAAGCTAACCTAAGAAAC 2152

Qy       5061 CCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCA 5120
Db        2153 CCACCTCTCATGGTTTGCTCTGTTGCCTATAGCCTTAGTAAGAATCCGAAACTCTCCCCA 2212

Qy       5121 AAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCT 5180
Db        2213 AAAAGCGGGACTCAGCCCATACGAAATGCTGTATGGACGGCCCTTCCTAACCAATGACCT 2272

Qy       5181 TGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 5239
Db        2273 TGTGCTTGACCTAGAGATGGCCAACTTAGTTGCAGATATC-CCTCCTTAGCCAAATATCA 2331

Qy       5240 ACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCA 5299
Db        2332 ACAAGTTCTTAAAACGTCACAGGGAACCTGTCCCTGAGAGGAGGGAAAGGAATTATTCCA 2391

Qy       5300 CCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCAT-CCCTAGATACAT 5358
Db        2392 ACCTGGTGACATGGTATTAGTGAAGTCCCTTCCCTCCAACTCCCCATCCCCTGGATACAT 2451

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
Db        2452 CCTGGGAAGGACCCTACTCAGTCATTTTATCTATCCCAACCGCGGTTAAAATGGCTGGAG 2511

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCA-AAGGAACCTGAA 5477
Db        2512 TAGAATCTTGGATACATCACATTCGAGTCAAACCCTAGATACTGCCACAAGGAACCTGAA 2571

Qy       5478 AATCCAGGAGACAA 5491
Db        2572 AATCCAGGAGACAA 2585
```

```
RESULT 23
ABN97948
ID   ABN97948 standard; DNA; 2055 BP.
XX
AC   ABN97948;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Human retroviral HERV-7q env coding sequence.
XX
KW   Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW   multiple sclerosis; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9967395-A1.
XX
PD   29-DEC-1999.
XX
PF   23-JUN-1999;   99WO-FR001513.
XX
PR   23-JUN-1998;   98FR-00007920.
```

XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX
PS Claim 3; Page 142-145; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 U; 0 Other;

Query Match 27.1%; Score 2032.5; DB 1; Length 2055;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 2025; Conservative 30; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db         61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
             |||||||||||||||||||||||||| ||||||||:|||||||||||||||||||||||:
Db        361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
             ||:||||:|||:|||||:|||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
             ||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||
Db        480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||:||||||
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
             ||||||||||||||||||||||||:|||||||||:|||||||||||||||||:||||||
Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
```

```
                :|||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
                |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
                |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
                |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245
                ||||||||||||||||
Db       2040 TAGGTTTTCCTGTTGA 2055
```

RESULT 24
AAV43199
ID AAV43199 standard; cDNA; 2304 BP.
XX
AC AAV43199;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus-1 pol gene fragment.
XX
KW Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 1. .2304
FT /*tag= a
FT /product= "MSRV pol protein fragment"
FT /note= "CDS does not contain a stop codon"
XX
PN WO9823755-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-IB001482.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX
DR WPI; 1998-322732/28.
DR P-PSDB; AAW71062.
XX
PT New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT treatment of, e.g. multiple sclerosis.
XX
PS Claim 1; Page 174-175; 286pp; English.
XX
CC The present sequence represents the multiple sclerosis (MS) associated
CC retrovirus-1 (MSRV-1) pol gene fragment. The invention provides complete
CC or partial genomic sequences of the MSRV-1 pol gene, gag gene and env
CC gene, and polypeptides encoded by these genes. The invention also
CC provides antibodies raised against the polypeptides. The genomic
CC sequences, polypeptides and antibodies are also claimed useful for
CC diagnosing infection by MS and rheumatoid arthritis-associated viruses,
CC and also for prevention and treatment of infection with these viruses
XX
SQ Sequence 2304 BP; 673 A; 561 C; 503 G; 566 T; 0 U; 1 Other;

Query Match 24.7%; Score 1851.9; DB 1; Length 2304;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2032; Conservative 119; Mismatches 150; Indels 29; Gaps 15;

```
Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
                |||||||||||||||||||||||||| |||||||| |||| ||||||||||||||||  |||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
                ||||| ||||||| |||||||||||| ||||||||  |   ||||||||||||||||  |
Db         59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
                | | ||||| |||||||||  | |||||||  ||||| |||||||||||||||||||||  |
Db        119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                  | |||  | || ||| | |||||||||  |||| ||:|| ||||||||||||| || |
Db        179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
                ||||||| ||||| |||||||||||||||||||||||||| |||||||||||||  ||||
Db        234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||||| ||||
Db        294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
                || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||| ||
Db        354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
                |||||||||||||||  || ||||||||||||||||||||||||||||||||||||||||
Db        413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
                | |||||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db        472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
                |||||||||||||||  |||||||||||||  |||||||||||||||  |  |||||||  |
Db        532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
                | | |        |||  |  | |  ||||||||  ||||||||||  ||||||||||||
Db        592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
                |||||||||||||  ||| ||||| ||| | | | ||||||||||||||  ||||||||||||
Db        644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
                ||| ||::||||| |||||||||||||||||||||||||||:||||||||||:|||||
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
                |||||  |||||||||||||||  |||||||||||||||||||||||||||||||||  ||
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
                |  ||||||:| |||||||||||||||||:|||||||||||||||||||||||||:|||
Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
```

```
                |||||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||||
Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
                |||||||:||| ||||||||||||||||||:| :|||::|||||||||||||||||||||
Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
                |||||||||| :|||||:| |||||||| |:|||||||||||||:|||||||||||:|:|:
Db       1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
                |||| :|||||||||||| || ||| |||||||||:||||:|||||||||||||||||:|
Db       1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
                |||||||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||
Db       1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
                ||||||||:|::||||||||||:|||:|||:||||||||||:|  |||| |||||||:
Db       1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
                |  ||||||||| :|||||:|||||:|||||| |||||||||||||||||||||| ||||
Db       1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
                |:|||||||||:|:||    :::|||||||||||||||||  ||  | | |||  ||||
Db       1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
                ||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: ||||| ||||
Db       1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
                |||||:||:||||||:|||||||||| |||:|| |||||:|||||||||||::|||
Db       1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
                ::||||||:||||||||:|||||||||||||||||||||:|:||||||||||:||||| |
Db       1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
                || ||||||||||||::|||||||||||||||||||||||||||||:||||||||||
Db       1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
                |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| ||||
Db       1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
                |||||||: ||||||||| |:|||||||||||||||||||: |||||||||| :|||| |
Db       1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
                |||||||||||||||||| || |:| |||||||||||||||||||||||||||||:||:|
Db       1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
                |||||||||||||:||||| |||||||||||||  |||||||||||||||:||||:|||||
Db       1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
                :|||||||||||:|||| |||||||||||||||||||||||||||||| | ||||||||||
Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
                |||||||||:|||||||||||||| ||||||||||||||||:|||||||||||| |||||
Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
                |||||||||||||||||:| ||:||||||||||||||||:|||:|:|||||||||||||||
```

```
Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
                ||||||:|||||||||:|||||||||| |||||||||:|:||||||:|||||||| ||||||
Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
                ||||||::|||||||||:|:||| |||||||||||||||||:|||:||||||||||: |
Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
                ||||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:|||||
Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

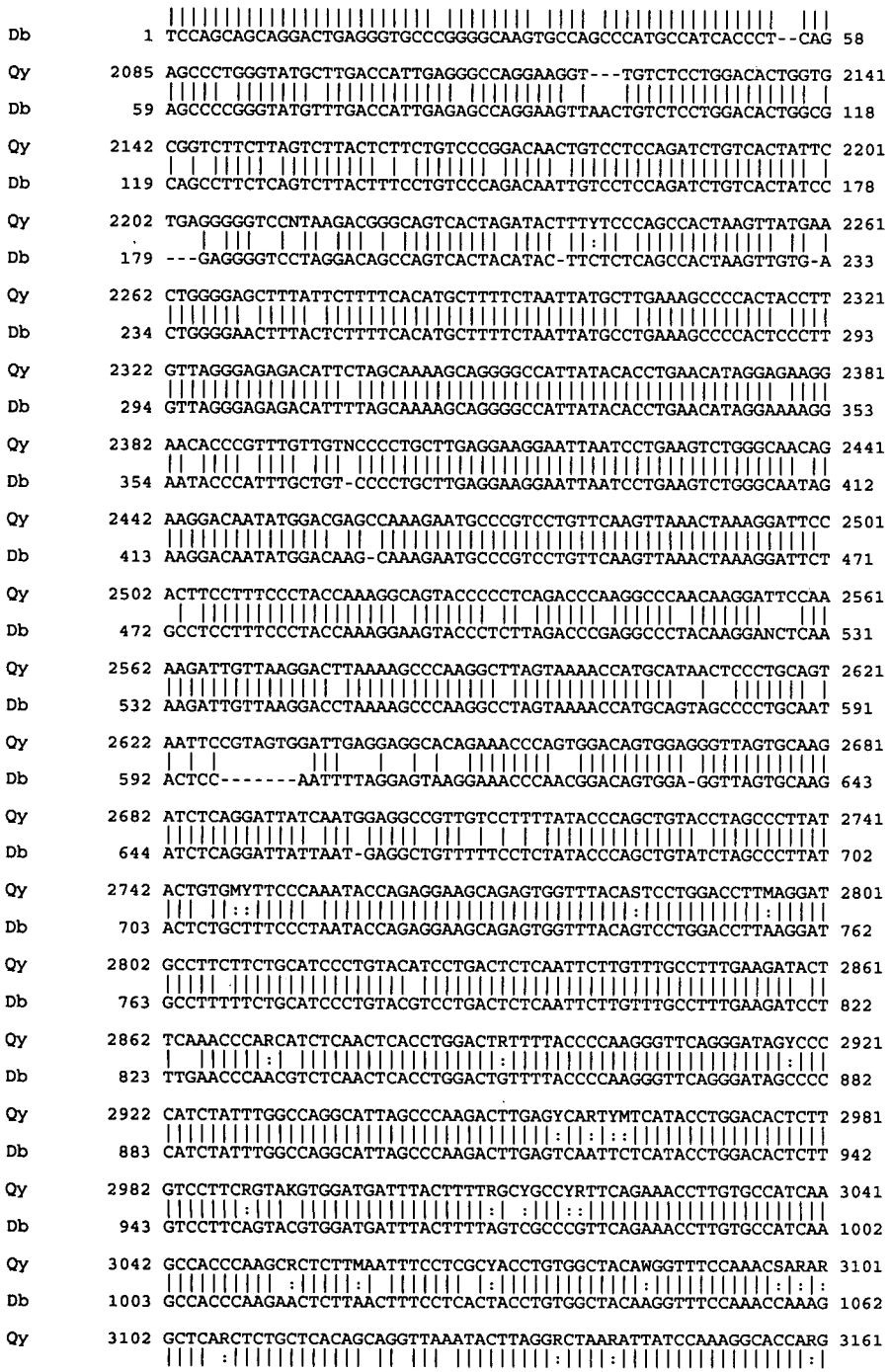
Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
                ||||||||: |||||||||  ||| | ||:|||||||||||:|||: ||: ||||||| :
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
                ||||||||||:|||| ||| |||||||||||| ||||||||||||||||
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304
```

RESULT 25
ADG14828
ID  ADG14828 standard; cDNA; 2304 BP.
XX
AC  ADG14828;
XX
DT  26-FEB-2004  (first entry)
XX
DE  MSRV associated cDNA #4.
XX
KW  ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS  Multiple sclerosis associated retrovirus.
XX
PN  US2003198647-A1.
XX
PD  23-OCT-2003.
XX
PF  03-APR-2002; 2002US-00114104.
XX
PR  26-NOV-1996;  96US-00756429.
PR  26-NOV-1997;  97US-00979847.
XX
PA  (INMR ) BIO MERIEUX.
XX
PI  Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI  Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR  WPI; 2004-032461/03.
XX
PT  New isolated nucleic acid and their fragments having the pol gene of a
PT  retrovirus, useful for diagnosing, preventing and/or treating multiple
PT  sclerosis and/or rheumatoid arthritis.
XX
PS  Claim 3; SEQ ID NO 87; 193pp; English.
XX
CC  The invention relates to an isolated nucleic acid which comprises the pol
CC  gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC  arthritis. The methods and compositions of the present invention are
CC  useful for diagnosing, preventing and/or treating multiple sclerosis
CC  and/or rheumatoid arthritis. The present sequence is used in the
CC  exemplification of the invention.
XX
SQ  Sequence 2304 BP; 673 A; 561 C; 503 G; 566 T; 0 U; 1 Other;

  Query Match            24.7%;  Score 1851.9;  DB 1;  Length 2304;
  Best Local Similarity  87.2%;  Pred. No. 0;
  Matches 2032;  Conservative  119;  Mismatches  150;  Indels   29;  Gaps   15;

```
Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
```


Db 1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58
Qy 2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
Db 59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118
Qy 2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
Db 119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178
Qy 2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
Db 179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233
Qy 2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
Db 234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293
Qy 2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
Db 294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353
Qy 2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
Db 354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412
Qy 2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
Db 413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471
Qy 2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
Db 472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531
Qy 2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
Db 532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591
Qy 2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
Db 592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643
Qy 2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
Db 644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702
Qy 2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
Db 703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762
Qy 2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
Db 763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822
Qy 2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
Db 823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882
Qy 2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
Db 883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942
Qy 2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
Db 943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002
Qy 3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
Db 1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062
Qy 3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
Db 1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122
Qy 3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
Db 1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182
Qy 3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
Db 1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241
Qy 3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
Db 1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301
Qy 3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
Db 1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356
Qy 3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
Db 1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415
Qy 3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
Db 1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475
Qy 3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
Db 1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535
Qy 3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
Db 1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595
Qy 3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
Db 1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655
Qy 3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
Db 1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715
Qy 3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
Db 1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775
Qy 3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
Db 1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835
Qy 3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
Db 1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894
Qy 3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
Db 1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954
Qy 4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
Db 1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014
Qy 4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
Db 2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074
Qy 4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
Db 2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134
Qy 4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
Db 2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

```
Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: |||||||||  ||| | ||:|||||||||||:|||: ||: ||||||| :
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
              |||||||||||:|||| ||| |||||||||||| ||||||||||||||||
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 26
AAX25658
ID   AAX25658 standard; cDNA to mRNA; 2006 BP.
XX
AC   AAX25658;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.Pi22.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 54-56; 106pp; French.
XX
CC   This sequence represents clone cl.Pi22 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 2006 BP; 625 A; 532 C; 373 G; 475 T; 0 U; 1 Other;

  Query Match             24.7%;  Score 1848.6;  DB 1;  Length 2006;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1949;  Conservative   56;  Mismatches    1;  Indels  206;  Gaps    3;

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
              ||:||||:|||||||||||||||:|||:|:|||||||||||||||||||:||||||||
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
              :|||||||||||||||||:|:|||||:|||||||| |||||||||||::|||||||
Db        121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
              |:|:|||||||||||||||||||||:|||:|||||||||||:||||||||||||||||||
Db        181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
              |||:|||||||:||||:|||||||||||||:||||||||:|||||||||||||||:||||||
Db        241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
              ||||||||||||||:|
Db        301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 ------------------------------------------------------------ 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 ------------------------------------------------------------ 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
              ||||||||||||||||||||
Db        317 ----------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
              |:|||:|||||||||||||:|||||||:|:|||||||||||||||:||||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
              ||||||||:|||||||:|||||||||||||||||:|:||:|||||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
```

```
                |||||||
```

```
Db        317 ------------------------------------------------------------ 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
                                                  ||||||||||||||||||||||
Db        317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
              |:|||:|||||||||||||:|||||||:|:|||||||||||||||:||||||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
              ||||||||:|||||||:|||||||||||||||||:|:||:|||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
              ||||||||||||||||||||||||||::||||||||||||||||||||||||:||||
Db       1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
              |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db       1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
              |||||||||:|||||||||||||||||||||||||:||:|||||:|||:||||||:||||||||
Db       1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
              ||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||
Db       1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006
```

```
RESULT 28
ADB84382
ID   ADB84382 standard; DNA; 2304 BP.
XX
AC   ADB84382;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 gene encoding reverse transcriptase/RNaseH.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
```

```
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI   Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84540.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 18; Fig 46; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 2304 BP; 673 A; 562 C; 503 G; 565 T; 0 U; 1 Other;

  Query Match             24.7%;  Score 1847.5;  DB 1;  Length 2304;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 2030;  Conservative  118;  Mismatches  153;  Indels   29;  Gaps   15;

Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             ||||||||||||||||||||||||| |||||||| |||| |||||||||| |||| |||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCGCCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| ||||||| |||||||||||| ||||||||| |   |||||||||||||||| |
Db         59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             | | ||||| ||||||||| | ||||||| ||||| | |||||||||||||||||||| |
Db         119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTATCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                | |||  | || ||| | ||||||||| |||| ||:|| |||||||||||||| || |
Db         179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
             ||||||| ||||| |||||||||||||||||||||||||||||||| |||||||||||| ||||
Db         234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||

Db        294 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
             || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
             ||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
              | |||||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db        472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
             |||||||||||||||| |||||||||||||| ||||||||||||||| | |||||||  |
Db        532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
             | | |        |||   | | |  |||||||| |||||||||  ||||||||||||
Db        592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
             |||||||||||| ||| ||||| ||| | | | ||||||||||||||| ||||||||||
Db        644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
             ||| ||::||||| ||||||||||||||||||||||||||||:||||||||||:|||||
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||| ||
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
             |  |||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:|||
Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
             |||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||||
Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
             |||||||:||| |||||||||||||||||:| :|||::|||||||||||||||||||||||
Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
             |||||||||| :|||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:
Db       1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
             |||| :|||||||||||| || ||| |||||||||:||||:|||||||||||||||||:|
Db       1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
             |||||||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||
Db       1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
             |||||||||::|::|||||||||:|||:|||:|||||||||||:|  |||| |||||||:
Db       1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
             |  |||||||||  :|||||:|||||:|||||| ||||||||||||||||||||||||||||
Db       1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
             |:|||||||||:|:||    :::|||||||||||||||||  ||  | | |||  ||||
Db       1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
             ||||:|||||||||  ||:|||||||:|||||||  ||||:||:|||:: ||||| ||||
Db       1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415
```

```
Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
Db       1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
Db       1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
Db       1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
Db       1596 TGTGTGGACATCTCATGATGTGAACCGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
Db       1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
Db       1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
Db       1776 TAACTGTGAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304
```

```
RESULT 29
AAD41225
ID   AAD41225 standard; cDNA; 2074 BP.
XX
AC   AAD41225;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Human EMBRY-2 cDNA.
XX
KW   Human; embryogenesis associated protein; AIDS; reproductive disorder;
KW   infertility; endometriosis; endometrial tumour; inflammatory disorder;
KW   autoimmune disorder; acquired immune deficiency syndrome; transgenic;
KW   ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
KW   EMBRY-2; allergy; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             64. .1692
FT                   /*tag= a
FT                   /product= "EMBRY-2 protein"
FT   sig_peptide     64. .123
FT                   /*tag= b
FT   mat_peptide     124. .1689
FT                   /*tag= c
FT                   /product= "Mature EMBRY-2 protein"
XX
PN   WO200248362-A2.
XX
PD   20-JUN-2002.
XX
PF   14-NOV-2001; 2001WO-US043956.
XX
PR   15-NOV-2000; 2000US-0249407P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Ramkumar J,  Arvizu C;
XX
DR   WPI; 2002-537629/57.
DR   P-PSDB; AAE25054.
XX
PT   New polypeptides of human embryogenesis associated proteins for screening
PT   modulators useful for treating or preventing disorders e.g.
PT   endometriosis, infertility, allergy, preeclampsia.
XX
PS   Claim 59; Page 96-97; 97pp; English.
XX
CC   The invention relates to human embryogenesis associated proteins (EMBRY)
CC   and nucleic acid molecules encoding such proteins. EMBRY sequences are
CC   useful for screening modulators useful for treating or preventing
CC   disorders associated with abnormal expression of EMBRY. The disorders
CC   treated include reproductive disorders such as infertility,
CC   endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
CC   disorder such as acquired immune deficiency syndrome (AIDS), allergies,
CC   contact dermatitis; disorders of the placenta such as preeclampsia,
CC   abruptio placentae etc. Sequences of the invention are also useful for
CC   analysing a proteome of a tissue or a cell type. EMBRY proteins are
CC   useful as immunogens for preparing antibodies. Polynucleotides of the
CC   invention are useful for creating knockin humanised animals or transgenic
CC   animals to model human diseases. They are also used in gene therapy. The
CC   present sequence is human EMBRY-2 cDNA
XX
SQ   Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;
```

```
  Query Match             24.2%;  Score 1812.4;  DB 1;  Length 2074;
  Best Local Similarity   91.7%;  Pred. No. 0;
  Matches 1904;  Conservative   33;  Mismatches  125;  Indels   14;  Gaps    3;

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
Db          1 GAGGATCTGTGCCTGCTCTTCAAGCAACAACCGTGAGGAAAGTAACTAAAATCGTAAAT- 59

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
Db         60 CCCCATGGCCCTCCCTTATTGTATTTTTCTCTTTACTGTTCTCTCACCACCTTTCAGTCT 119

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
Db        120 CACTGCACCCTCTCCATGCCACTGTAGGACCAGTAGCTCCCCTTACCAAGCGTTTCTATG 179

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
```

```
Db        180 GAGAATGCGGCGTCCCAGACATATTGATGCCCCATCGTATAGGAGTTTATCTAAGGGAAA 239
Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
Db        240 CCCCGCCTTCACCGCCCACACCCATATGCCCCACAACTGCTATAACTCTGCCACTCTTTG 299

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
Db        300 TATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGG 359

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
Db        360 ACTTGGAGCCACTATCTGTTGGACTTACTTCACCCATACCGGTATGTCTGATGGGGGTGG 419

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
Db        420 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGGAAGTAATCTCCCAACTCACCCGGGT 479

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
Db        480 ACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 539

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
Db        540 TACCCATACTCACCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 599

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTC 6176
Db        600 GGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTC 659

Qy       6177 AATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTT 6236
Db        660 AGTCCCTGTACCTGAACAATGGAATAACTTCAGCACAGAAATAAACACCACTTCCATTTT 719

Qy       6237 AGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAA 6296
Db        720 AGTAGGACCTCTTGTTTCCAATCTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAA 779

Qy       6297 ATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAC 6356
Db        780 ATTTAGCAATACTATAGCAACAACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCAC 839

Qy       6357 ACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTG 6416
Db        840 ACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTG 899

Qy       6417 TTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCAT 6476
Db        900 TTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCAT 959

Qy       6477 CTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACC 6536
Db        960 CTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACC 1019

Qy       6537 CATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGG 6596
Db       1020 CATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAG 1079

Qy       6597 TATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGA 6656
Db       1080 TATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGA 1139

Qy       6657 ACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGT 6716
Db       1140 ACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGT 1199

Qy       6717 CCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATT 6776
Db       1200 CCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATT 1259

Qy       6777 TTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRA 6836
Db       1260 TTTAGGAGAAGAATGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAA 1319

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
Db       1320 AGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACACTGGGG 1379

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
Db       1380 CCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAAT 1439

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
Db       1440 ATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 1499

Qy       7017 AATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
Db       1500 AATTGAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAGTCCATGAC 1559

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
Db       1560 TAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACAT 1619

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
Db       1620 CGAAGGCACTCCTCCCAAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGC 1679

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
Db       1680 AGGAAGCAGTTAGAGCGGTCGTCAGTCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTG 1739

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
Db       1740 AGAGGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTAAGAATCCCTAAG 1799

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
Db       1800 CCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAC 1859

Qy       7365 CTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATA 7424
Db       1860 CCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATA 1919

Qy       7425 GCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAA 7484
Db       1920 GCCAATCATCTATTGCCTGAGAGCACAGTGGGAGGGACAAGGATTGCAATATAAACCCAG 1979

Qy       7485 GTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGT 7544
Db       1980 GCATTCGAGCCAGC-ANAGCAACCGCCTTTGGGTCCCTTCCCCTTGTATGGGAGCTCTGT 2038

Qy       7545 TTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
Db       2039 TTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 2074
```

```
RESULT 30
AAA63826
ID   AAA63826 standard; DNA; 2030 BP.
XX
AC   AAA63826;
XX
DT   06-AUG-2003  (revised)
DT   04-DEC-2000  (first entry)
XX
DE   Nucleotide sequence of the MSRV-1 3' env and LTR regions.
XX
KW   MSRV-1; pol region; long terminal repeat; LTR; RU5 region; retrovirus;
KW   ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1629
FT                   /*tag=  a
```

```
FT                   /note= "Contains one termination codon"
FT   sig_peptide     1. .81
FT                   /*tag=  b
FT   CAAT_signal     1800. .1807
FT                   /*tag=  c
FT   CAAT_signal     1858. .1864
FT                   /*tag=  d
FT   TATA_signal     1906. .1911
FT                   /*tag=  e
FT   polyA_signal    1996. .2002
FT                   /*tag=  f
XX
PN   WO200047745-A1.
XX
PD   17-AUG-2000.
XX
PF   15-FEB-2000; 2000WO-IB000159.
XX
PR   15-FEB-1999;   99EP-00420041.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Perron H,  Komurian-Pradel F;
XX
DR   WPI; 2000-506097/45.
DR   P-PSDB; AAB08195.
XX
PT   Nucleotide fragment of LTR-RU5 region from Multiple Sclerosis retrovirus
PT   (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
PT   sample.
XX
PS   Disclosure; Fig 2; 23pp; English.
XX
CC   The present sequence represents the nucleotide sequence corresponding to
CC   the 3' env region and long terminal repeat sequences from clone CL6 of
CC   Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
CC   long terminal repeat (LTR)-RU5 region which encodes the expression of a
CC   MSRV-1 protein. This is unusual for LTRs, in particular in the RU5
CC   region. The sequence includes CAAT and TATA signals which are present in
CC   the U3 and R regions and are not directed towards the CDS indicated in
CC   the features table. Probes and antibodies to the MSRV-1 retrovirus
CC   protein and encoding polynucleotide sequences are used to detect the
CC   presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
CC   2003 to correct OS field.)
XX
SQ   Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;

  Query Match             23.8%;  Score 1780.4;  DB 1;  Length 2030;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches 1848;  Conservative   33;  Mismatches  119;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440
```

```
Qy       7021 GAAGCTGTRAAAC----------TACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCAGATGGAGTCCATGACTAAG 1500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
Db       1501 ATCCACCGTGGACCCCTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAA 1560

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
Db       1561 GGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGA 1620

Qy       7189 AGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAT 7248
Db       1621 AGCAGTTAGAGCGGTCATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAG 1680

Qy       7249 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTA 7308
Db       1681 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTA 1740

Qy       7309 GSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGA 7368
Db       1741 GCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGA 1800

Qy       7369 CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCA 7428
Db       1801 CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCA 1860

Qy       7429 ATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYT 7488
Db       1861 ATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCAT 1920

Qy       7489 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTC 7548
Db       1921 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTC 1980

Qy       7549 ATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
Db       1981 ACTCTATTTCACTCTATTAAATCTTGCAACTG 2012
```

```
RESULT 31
AAH20070
ID   AAH20070 standard; DNA; 1617 BP.
XX
AC   AAH20070;
XX
DT   08-AUG-2001  (first entry)
XX
DE   HERV-W envelope protein G coding sequence.
XX
KW   Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW   envelope protein; multiple sclerosis-related superantigen; vaccine;
KW   surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW   antisense-therapy; autoimmune disorder; ds.
XX
OS   Human endogenous retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1617
FT                   /*tag=  a
FT                   /product= "HERV-W envelope protein G"
XX
PN   WO200131021-A1.
XX
PD   03-MAY-2001.
XX
PF   30-OCT-2000; 2000WO-EP010659.
XX
PR   28-OCT-1999;   99EP-00402690.
XX
PA   (UYGE-) UNIV GENEVE.
XX
PI   Conrad B,  Mach B;
XX
DR   WPI; 2001-316336/33.
DR   P-PSDB; AAB75138.
XX
PT   New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT   activity useful for diagnosing and treating multiple sclerosis.
XX
PS   Claim 13; Fig 10; 94pp; English.
XX
CC   On the basis of the PBS t-RNA motif used for the classification of human
CC   endogenous retrovirus (HERVs) the full length endogenous provirus which
CC   was been located on the long arm of human chromosome 7 (7q21-22) has been
CC   designated HERV-W. The present invention describes proteins or peptides
CC   (I) having superantigen (SAg) activity comprising the ENV protein (ENV)
CC   of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC   have neuroprotective activity, and can be used in: vaccines; antisense-
CC   therapy; and HERV-W SAg activity-inhibitors. (I) and encoding DNA/RNA are
CC   useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC   disorders. (I) are also useful for identifying substances (and optionally
CC   recovering) capable of binding to a retroviral superantigen associated
CC   with MS, substances capable of blocking SAg activity and substances
CC   capable of blocking transcription or translation of HERV-W retroviral
CC   superantigen. A protein or peptide derived from (I), modified to be
CC   devoid of SAg activity and being capable of generating an immune response
CC   against HERV-W retroviral SAg is useful in therapy. Nucleic acid
CC   molecules encoding (I) are useful as vaccines against MS. Substances
CC   capable of blocking SAg activity, capable of binding to a retroviral
CC   superantigen associated with MS, or capable of blocking transcription or
CC   translation of HERV-W retroviral superantigen for use in treating or
CC   preventing MS, obtained using (I) are useful for the treatment and
CC   prevention of MS. (I) and nucleic acids encoding them are useful for
CC   diagnosing autoimmune disease. The present sequence encodes the
CC   specifically claimed envelope protein of HERV-W designated G
XX
SQ   Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;
```

```
  Query Match             21.5%;  Score 1607;  DB 1;  Length 1617;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1592;  Conservative   25;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480
```

```
Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:|||||||||||||||||||:|||||||||||:|||||||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db       1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db       1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 1617
```

RESULT 32
AED10942
ID AED10942 standard; DNA; 1617 BP.
XX
AC AED10942;
XX
DT 15-DEC-2005 (first entry)
XX
DE HERV envelope protein DNA SEQ ID NO 73.
XX
KW virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW screening; immune modulation; graft rejection; allergy;
KW autoimmune disease; cancer; viral infection; envelope; env; gene; ds.
XX
OS Human endogenous retrovirus W.
XX
PN WO2005095442-A1.
XX
PD 13-OCT-2005.
XX
PF 30-MAR-2005; 2005WO-EP003339.
XX
PR 30-MAR-2004; 2004EP-00290838.
XX
PA (INSR ) INST ROUSSY GUSTAVE.
PA (CNRS ) CENT NAT RECH SCI.
PA (UYPA-) UNIV PARIS SUD.
XX
PI Renard M, Mangeney M, Heidmann T;
XX
DR WPI; 2005-703311/72.
DR P-PSDB; AED10943.
XX
PT New polypeptide capable of modulating the immunosuppressive properties of
PT a viral protein or its fragment, useful for the manufacture of a vaccine
PT for treating or preventing viral infections or cancer.
XX
PS Disclosure; SEQ ID NO 73; 262pp; English.
XX
CC The invention describes a polypeptide of 7-20 amino acid residues, which
CC is capable of modulating the immunosuppressive properties of a viral
CC protein or its fragment, against the host in which it is expressed
CC (immunosuppression-modulatory sequence) when it substitutes the
CC homologous sequence of the viral protein or its fragment, the polypeptide
CC comprising the minimum following consensus amino acid sequence of (I).
CC Also described are: a mutated ENV protein resulting from the mutation of
CC a wild type ENV protein essentially carrying the following sequence: NY 1
CC Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC or prokaryotic expression vector comprising the nucleic acid as well as
CC the elements necessary for the expression of the nucleic acid; a
CC recombinant cell comprising the nucleic acid or the eukaryotic or
CC prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC comprising as active substance: at least one polypeptide, mutated ENV
CC protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC expression vector, or recombinant cell; in association with a
CC pharmaceutically acceptable carrier; and antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides, directed against
CC mutated ENV proteins, provided that the antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides do not bind to the
CC corresponding wild type ENV proteins. The first mutation of a first amino
CC acid and optionally of a second mutation of a second amino acid in a wild
CC type viral envelope (ENV) protein is useful for manufacturing a mutated
CC ENV protein having a modified immunosuppressive activity with respect to
CC the wild type ENV protein. The protein is useful for the manufacture of a
CC medicament or a vaccine intended for the prevention and/or the treatment
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC requiring an inhibition of the immune system, such as autoimmune

```
CC   diseases, allergies or graft rejections. The protein is also useful for
CC   screening compounds liable to modulate the immunosuppressive activity of
CC   viruses or tumor cells, where the compounds to screen are peptides, in
CC   particular peptides comprising from 5-30 amino acids, such as peptides
CC   originating from combinatorial peptide libraries. This sequence encodes
CC   HERV envelope protein.
XX
SQ   Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;

  Query Match             21.5%;  Score 1607;  DB 1;  Length 1617;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1592;  Conservative   25;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||:||||||||||:||||||||||||||||||||||||||||
Db        1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db        1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 1617


RESULT 33
AED10984
ID   AED10984 standard; DNA; 1614 BP.
XX
AC   AED10984;
XX
DT   15-DEC-2005  (first entry)
XX
DE   HERV R393E/Q envelope protein DNA SEQ ID NO 115.
XX
KW   virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW   screening; immune modulation; graft rejection; allergy;
KW   autoimmune disease; cancer; viral infection; envelope; env; gene; ds;
KW   mutant.
XX
OS   Human endogenous retrovirus W.
XX
PN   WO2005095442-A1.
XX
PD   13-OCT-2005.
XX
PF   30-MAR-2005; 2005WO-EP003339.
XX
PR   30-MAR-2004; 2004EP-00290838.
```

XX
PA (INSR ) INST ROUSSY GUSTAVE.
PA (CNRS ) CENT NAT RECH SCI.
PA (UYPA-) UNIV PARIS SUD.
XX
PI Renard M, Mangeney M, Heidmann T;
XX
DR WPI; 2005-703311/72.
DR P-PSDB; AED10985.
XX
PT New polypeptide capable of modulating the immunosuppressive properties of
PT a viral protein or its fragment, useful for the manufacture of a vaccine
PT for treating or preventing viral infections or cancer.
XX
PS Claim 41; SEQ ID NO 115; 262pp; English.
XX
CC The invention describes a polypeptide of 7-20 amino acid residues, which
CC is capable of modulating the immunosuppressive properties of a viral
CC protein or its fragment, against the host in which it is expressed
CC (immunosuppression-modulatory sequence) when it substitutes the
CC homologous sequence of the viral protein or its fragment, the polypeptide
CC comprising the minimum following consensus amino acid sequence of (I).
CC Also described are: a mutated ENV protein resulting from the mutation of
CC a wild type ENV protein essentially carrying the following sequence: NY 1
CC Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC or prokaryotic expression vector comprising the nucleic acid as well as
CC the elements necessary for the expression of the nucleic acid; a
CC recombinant cell comprising the nucleic acid or the eukaryotic or
CC prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC comprising as active substance: at least one polypeptide, mutated ENV
CC protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC expression vector, or recombinant cell; in association with a
CC pharmaceutically acceptable carrier; and antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides, directed against
CC mutated ENV proteins, provided that the antibodies or their fragments,
CC scFv polypeptides, aptamers, or binding peptides do not bind to the
CC corresponding wild type ENV proteins. The first mutation of a first amino
CC acid and optionally of a second mutation of a second amino acid in a wild
CC type viral envelope (ENV) protein is useful for manufacturing a mutated
CC ENV protein having a modified immunosuppressive activity with respect to
CC the wild type ENV protein. The protein is useful for the manufacture of a
CC medicament or a vaccine intended for the prevention and/or the treatment
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC requiring an inhibition of the immune system, such as autoimmune
CC diseases, allergies or graft rejections. The protein is also useful for
CC screening compounds liable to modulate the immunosuppressive activity of
CC viruses or tumor cells, where the compounds to screen are peptides, in
CC particular peptides comprising from 5-30 amino acids, such as peptides
CC originating from combinatorial peptide libraries. This sequence encodes
CC HERV envelope protein.
XX
SQ Sequence 1614 BP; 439 A; 452 C; 294 G; 426 T; 0 U; 3 Other;

Query Match 21.4%; Score 1601; DB 1; Length 1614;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1586; Conservative 25; Mismatches 3; Indels 0; Gaps 0;

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||||:||||||||||||:   ||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAANNNGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:||||||||||||||||:|||||||:||||||||||||||||||||||||||||||
Db       1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320
```

```
Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db       1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 1614


RESULT 34
AED10986
ID   AED10986 standard; DNA; 1614 BP.
XX
AC   AED10986;
XX
DT   15-DEC-2005  (first entry)
XX
DE   HERV R393E/Q F399A envelope protein DNA SEQ ID NO 117.
XX
KW   virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
KW   screening; immune modulation; graft rejection; allergy;
KW   autoimmune disease; cancer; viral infection; envelope; env; gene; ds;
KW   mutant.
XX
OS   Human endogenous retrovirus W.
XX
PN   WO2005095442-A1.
XX
PD   13-OCT-2005.
XX
PF   30-MAR-2005; 2005WO-EP003339.
XX
PR   30-MAR-2004; 2004EP-00290838.
XX
PA   (INSR ) INST ROUSSY GUSTAVE.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (UYPA-) UNIV PARIS SUD.
XX
PI   Renard M,  Mangeney M,  Heidmann T;
XX
DR   WPI; 2005-703311/72.
DR   P-PSDB; AED10987.
XX
PT   New polypeptide capable of modulating the immunosuppressive properties of
PT   a viral protein or its fragment, useful for the manufacture of a vaccine
PT   for treating or preventing viral infections or cancer.
XX
PS   Claim 41; SEQ ID NO 117; 262pp; English.
XX
CC   The invention describes a polypeptide of 7-20 amino acid residues, which
CC   is capable of modulating the immunosuppressive properties of a viral
CC   protein or its fragment, against the host in which it is expressed
CC   (immunosuppression-modulatory sequence) when it substitutes the
CC   homologous sequence of the viral protein or its fragment, the polypeptide
CC   comprising the minimum following consensus amino acid sequence of (I).
CC   Also described are: a mutated ENV protein resulting from the mutation of
CC   a wild type ENV protein essentially carrying the following sequence: NY 1
CC   Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic
CC   acid coding for the polypeptide or the mutated ENV protein; an eukaryotic
CC   or prokaryotic expression vector comprising the nucleic acid as well as
CC   the elements necessary for the expression of the nucleic acid; a
CC   recombinant cell comprising the nucleic acid or the eukaryotic or
CC   prokaryotic expression vector; a pharmaceutical or a vaccine composition
CC   comprising as active substance: at least one polypeptide, mutated ENV
CC   protein or its fragments, nucleic acid, prokaryotic or eukaryotic
CC   expression vector, or recombinant cell; in association with a
CC   pharmaceutically acceptable carrier; and antibodies or their fragments,
CC   scFv polypeptides, aptamers, or binding peptides, directed against
CC   mutated ENV proteins, provided that the antibodies or their fragments,
CC   scFv polypeptides, aptamers, or binding peptides do not bind to the
CC   corresponding wild type ENV proteins. The first mutation of a first amino
CC   acid and optionally of a second mutation of a second amino acid in a wild
CC   type viral envelope (ENV) protein is useful for manufacturing a mutated
CC   ENV protein having a modified immunosuppressive activity with respect to
CC   the wild type ENV protein. The protein is useful for the manufacture of a
CC   medicament or a vaccine intended for the prevention and/or the treatment
CC   of viral diseases, such as HTLV or FeLV infections; cancer; pathologies
CC   requiring an inhibition of the immune system, such as autoimmune
CC   diseases, allergies or graft rejections. The protein is also useful for
CC   screening compounds liable to modulate the immunosuppressive activity of
CC   viruses or tumor cells, where the compounds to screen are peptides, in
CC   particular peptides comprising from 5-30 amino acids, such as peptides
CC   originating from combinatorial peptide libraries. This sequence encodes
CC   HERV envelope protein.
XX
SQ   Sequence 1614 BP; 439 A; 452 C; 294 G; 423 T; 0 U; 6 Other;

  Query Match             21.3%;  Score 1598;  DB 1;  Length 1614;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches 1583;  Conservative   25;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660
```

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:|||||||||||||||||:|||||||||||:   |||||||||||||||   |||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAANNNGGGGGAACCTGTTTANNNTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db       1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||||||:||||||:|||||||||||||||||||||||||||||||
Db       1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db       1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 7194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGT 1614


RESULT 35
AAI14608
ID   AAI14608 standard; DNA; 1894 BP.
XX
AC   AAI14608;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #4541 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 4541; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match            20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity  90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||          |||||  ||||||  ||  ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347
```

```
Qy      6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db       348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy      6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy      6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||  ||||  |||||||:|||||||||||||||||||||||||
Db       468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy      6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db       525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy      6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy      6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db       645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy      6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db       705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy      6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db       765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy      6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy      6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||| |||||||  ||  ||||||||||||||||||||||||||||||||
Db       885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy      6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:||||||||| ||:|| ||||
Db       945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy      6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| |||||||||| ||||||||||||||||||||||| |||
Db      1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy      6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db      1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy      6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db      1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy      6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| |||||||||||||||||||||||||||:|||||||||||| |
Db      1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy      7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||| ||:||||||           |||||||||||| |||
Db      1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy      7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| |||||||| | ||||||||||||||||||:|||||||| | ||||
Db      1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy      7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||
Db      1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy      7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||||||||| ||| || |||||||||||||||||| ||
Db      1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy      7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||||| || || |||||||||| |  |:|
Db      1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy      7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||||| | |||||||||||||| |||||||||
Db      1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy      7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             |||||||:|||| ||  |||||||          ||||||||||||||||||||||||||||
Db      1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy      7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db      1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy      7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db      1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy      7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          ||||||||||||||||||||:||||:
Db      1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 36
ABA56337
ID  ABA56337 standard; DNA; 1894 BP.
XX
AC  ABA56337;
XX
DT  01-FEB-2002  (first entry)
XX
DE  Human foetal liver single exon nucleic acid probe #4642.
XX
KW  Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157277-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US000669.
XX
PR  04-FEB-2000; 2000US-0180312P.
PR  26-MAY-2000; 2000US-0207456P.
PR  30-JUN-2000; 2000US-00608408.
PR  03-AUG-2000; 2000US-00632366.
PR  21-SEP-2000; 2000US-0234687P.
PR  27-SEP-2000; 2000US-0236359P.
PR  04-OCT-2000; 2000GB-00024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR  WPI; 2001-483447/52.
XX
PT  Human genome-derived single exon nucleic acid probes useful for analyzing
PT  gene expression in human fetal liver.
XX
PS  Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX
CC  The invention relates to a single exon nucleic acid probe for measuring
CC  human gene expression in a sample derived from human foetal liver. The
CC  single exon nucleic acid probes may be used for predicting, measuring and
CC  displaying gene expression in samples derived from human fetal liver. The
CC  present sequence is a single exon nucleic acid probe of the invention.

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match           20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity  90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||              ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||| ||||   ||||||||:|||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:||||||||| ||:|| ||||
Db         945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| ||||||||| |||||||||||||||||||||||| |||
Db        1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:||||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db        1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||||:|||||||||||||| |
Db        1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||||| ||:|||||| |||||||||||| |||
Db        1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| |||||||| | |||||||||||||||||:|||||| | ||||
Db        1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db        1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||||||||| ||| || |||||||||||||| ||
Db        1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db        1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||| | |||||||||||||| |||||||||
Db        1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||          ||||||||||||||||||||||||||||
Db        1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db        1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db        1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          ||||||||||||||||||||:||||:
Db        1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

```
RESULT 37
AAI35980
ID   AAI35980 standard; DNA; 1894 BP.
XX
AC   AAI35980;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #4666 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
```

```
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 4666; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||          |||||  ||||||  ||  ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||  |||||||||   ||||  ||||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             ||||  |||  ||||||||  |||||  ||  |||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             ||||||||||||  |||||||||  |||||||||||||||||||||||||||||  |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||| ||||  |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :||||||  ||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             |||||||  |||||||||||  ||||   ||||||||:|||||||||||||||||||||||||
Db         468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||||| | |||||  ||||||||||||||||||||||||
Db         525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db         585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||||  ||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db         645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:|||  |||||||||||||||||||||||||||||  | ||||||||  ||||||||||||
Db         705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||  ||||||||||||||||||  | | ||| ||
Db         765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||  |||||||  ||  |||||||||||||||||||||||||||||||||
Db         885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:|||||||||  ||:||  ||||
Db         945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             ||||||  ||||||||||  ||||  |||||||||  ||||||||||||||||||||||  |||
Db        1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||||  ||  |||||||  ||||||:||||  |:|||||  | |||||
Db        1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||||  ||||||||||||||||||||||||||:|||||||||||  |
Db        1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051
             ||||||  ||||||||||||||||  ||:||||||          ||||||||||| |||
Db        1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             |||||||| ||  ||| |||||||  | ||||||||||||||||||:|||||||  | ||||
Db        1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||||||  ||||||||||||  ||||||||||||||||||
Db        1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||  |||||||||||||||||||  ||  ||  ||||||||||  |  |:|
Db        1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
```

```
                ||||||||:|||| ||||:|||||||||||||| | ||||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
                ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
                ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
                |||||||||||||||||||  |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                ||||||||||||||||||||          |||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 38
ABA45822
ID   ABA45822 standard; DNA; 1894 BP.
XX
AC   ABA45822;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #4517.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
                | |:|||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
                |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
                |||||||||||||||||||              ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
                |||||||||||||||||||||||||||||| ||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
                |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
                |||||||||||| ||||||||| |||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
                |||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
                :||||||  |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
                ||||||| |||||||||| ||||   ||||||||:|||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
                |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
                |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
                |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
                ||||:||| |||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
                |||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884
```

```
Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 39
ABA25978
ID ABA25978 standard; DNA; 1894 BP.
XX
AC ABA25978;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4444 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 1; SEQ ID NO 4444; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

```
  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
```

```
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||| |||| |||||||:||||||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  |||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             ||||||||||||||||||:||||||||||||||||||:|||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| |||||||| |||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:||||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| |||||||||||||||||||||:|||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||||| ||:|||||              |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| ||||||| | ||||||||||||||||||:||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||| ||| || ||||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||||| | |||||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||          |||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| |||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||  |||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 40
AAK30018
ID  AAK30018 standard; DNA; 1894 BP.
XX
AC  AAK30018;
XX
DT  06-NOV-2001 (first entry)
XX
DE  Human bone marrow expressed single exon probe SEQ ID NO: 4575.
XX
KW  Human; bone marrow expressed exon; gene expression analysis; probe;
KW  microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157276-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US000668.
XX
PR  04-FEB-2000; 2000US-0180312P.
PR  26-MAY-2000; 2000US-0207456P.
PR  30-JUN-2000; 2000US-00608408.
PR  03-AUG-2000; 2000US-00632366.
PR  21-SEP-2000; 2000US-0234687P.
PR  27-SEP-2000; 2000US-0236359P.
PR  04-OCT-2000; 2000GB-00024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-488900/53.
XX
PT  Human genome-derived single exon nucleic acid probes useful for analyzing
PT  gene expression in human bone marrow.
XX
PS  Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in the human
CC  bone marrow. They can be used to measure gene expression in bone marrow

```
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
Db         168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
Db         228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
Db         288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
Db         348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
Db         408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
Db         468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
Db         525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
Db         585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
Db         645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
Db         705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
Db         765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
Db         825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
Db         885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
Db         945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
Db        1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
Db        1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
Db        1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
Db        1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051
Db        1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
Db        1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
Db        1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
Db        1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
Db        1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
Db        1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
Db        1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
Db        1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
Db        1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
Db        1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834

RESULT 41
AAK04516
ID   AAK04516 standard; DNA; 1894 BP.
XX
AC   AAK04516;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 4507.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
```

```
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 4507; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223

Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
```

```
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             |||||||||:|||| ||||:|||||||||||| | |||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| |||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||          ||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 42
ABS29670
ID   ABS29670 standard; DNA; 1894 BP.
XX
AC   ABS29670;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 4660.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 4660; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||||             ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA-------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||||||| |||||||||  ||||  ||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :||||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||| ||||    ||||||||:|||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
```

```
                |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
                |||||||||||||||||:||||||||||||||||||:|||||||||  ||:||  ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
                ||||||  |||||||||||  ||||  |||||||||  ||||||||||||||||||||||  |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
                |  |||||||||:|||||||||  ||  |||||||  ||||||:||||  |:|||||  |  |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
                ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
                ||||||||||||||||  ||||||||||||||||||||||||:|||||||||||||  |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051
                ||||||  ||||||||||||||  ||:||||||          ||||||||||||  |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
                |||||||||  ||  |||  |||||||  |  ||||||||||||||||||:|||||||  |  ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
                |||||||||||||||||||||||||||||  ||||||||||||  ||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
                |||||||||||||||||||||||||||||  |||  ||  ||||||||||||||||||||  ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
                ||||||||||||||||  ||||||||||||||||||||  ||  ||  ||||||||||  |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
                ||||||||:||||  ||||:|||||||||||||  |  ||||||||||||||||  |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
                ||||||:||||  ||  |||||||          |||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
                |||  |||||||  ||||||||||||||||||||:|  ||||:|||  ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
                ||||||||||||||||||||  |  ||||||||  ||  ||  ||  ||||  ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                ||||||||||||||||||||          ||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 43
AAI04422
ID   AAI04422 standard; DNA; 1894 BP.
XX
AC   AAI04422;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #4413 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 4413; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
                | |:|||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||  |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
                ||||||||||||||||||||              |||||  |||||||  ||  ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
                |||||||||||||||||||||||||||||||  |||||||||  ||||  ||||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
                ||||  |||  |||||||||  |||||  ||  ||||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
                ||||||||||||  ||||||||||  ||||||||||||||||||||||||||||||  |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347
```

```
Qy      6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             ||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db       348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy      6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy      6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||| |||| ||||||||:|||||||||||||||||||||||||
Db       468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy      6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db       525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy      6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy      6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db       645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy      6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db       705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy      6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db       765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy      6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy      6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||
Db       885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy      6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:||||||||| ||:|| ||||
Db       945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy      6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| |||||||||| ||||||||||||||||||||| |||
Db      1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy      6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:||||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db      1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy      6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db      1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy      6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||||||:||||||||||||| |
Db      1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy      7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||| ||:|||||            |||||||||||| |||
Db      1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy      7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| |||||||| | ||||||||||||||||||:||||||| | ||||
Db      1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy      7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db      1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy      7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||||| ||| || |||||||||||||||||| ||
Db      1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy      7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||||| || || |||||||||| |  |:|
Db      1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy      7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||||| | |||||||||||||| |||||||||
Db      1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy      7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         ||||||||||||||||||||||||||||
Db      1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy      7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db      1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy      7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| | |||||||| || || || |||| ||||||||||||
Db      1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy      7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          ||||||||||||||||||||:||||:
Db      1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 44
ABS04589
ID  ABS04589 standard; DNA; 1894 BP.
XX
AC  ABS04589;
XX
DT  19-AUG-2002  (first entry)
XX
DE  Human genome-derived single exon probe from lung SEQ ID No 4580.
XX
KW  Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW  chronic obstructive pulmonary disease; interstitial lung disease;
KW  familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW  tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW  Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW  pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW  pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW  primary ciliary dyskinesis; pulmonary hypertension;
KW  hyaline membrane disease.
XX
OS  Homo sapiens.
XX
PN  WO200186003-A2.
XX
PD  15-NOV-2001.
XX
PF  30-JAN-2001; 2001WO-US000665.
XX
PR  04-FEB-2000; 2000US-0180312P.
PR  26-MAY-2000; 2000US-0207456P.
PR  30-JUN-2000; 2000US-00608408.
PR  03-AUG-2000; 2000US-00632366.
PR  21-SEP-2000; 2000US-0234687P.
PR  27-SEP-2000; 2000US-0236359P.
PR  04-OCT-2000; 2000GB-00024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR  WPI; 2002-114183/15.
XX
PT  Spatially-addressable set of single exon nucleic acid probes, used to
PT  measure gene expression in human lung samples.

```
XX
PS   Claim 1; SEQ ID NO 4580; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
              | |:||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
              ||||||||||||||||||||             ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
              |||||||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
              |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
              |||||||||||| |||||||||| ||||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
              :|||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
              ||||||| |||||||||||| |||| |||||||:|||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
              |||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
              |||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
              ||||:||| |||||||||||||||||||||||| | |||||||| ||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
              |||||||||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
              |||| |||||||| ||||||| ||  |||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
              |||||||||||||||||:|||||||||||||||:|||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
              |||||| |||||||||| |||| |||||||| |||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
              | |||||||||:|||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
              |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
              |||||| |||||||||||||| ||:|||||| |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
              ||||||||| || ||| |||||||| | ||||||||||||||||:|||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
              |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
              |||||||||||||||||||||||||||||| ||| || |||||||||||||||||| ||
```

```
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             |||||||:|||| |||||:|||||||||||| | |||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          |||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 45
ABN97931/c
ID   ABN97931 standard; DNA; 1799 BP.
XX
AC   ABN97931;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Human retroviral sequence HE3.
XX
KW   Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW   multiple sclerosis; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9967395-A1.
XX
PD   29-DEC-1999.
XX
PF   23-JUN-1999;   99WO-FR001513.
XX
PR   23-JUN-1998;   98FR-00007920.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Alliel PM,  Perin J,  Rieger F;
XX
DR   WPI; 2000-160587/14.
XX
PT   New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT   for diagnosis, treatment and prevention of autoimmune and neurological
PT   diseases.
XX
PS   Claim 3; Page 134; 225pp; French.
XX
CC   The present invention relates to new nucleic acid sequences of human
CC   endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC   Regulatory elements associated with HERV-7q may alter expression of other
CC   genes (even remote genes) on the same chromosome, inducing immunological
CC   and/or neurological changes (which may be pathological or protective/
CC   curative). HERV-7q peptides can be used to improve efficiency of the
CC   immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC   sequences can be used in immunogenic or vaccinating compositions, for
CC   protection against autoimmune diseases, particularly multiple sclerosis.
CC   The peptides may also be used (by sequence comparison) to detect/identify
CC   endogenous retroviruses that are abnormally expressed in cancer,
CC   neuropathologies or other autoimmune diseases. The present sequence was
CC   used to illustrate the invention
XX
SQ   Sequence 1799 BP; 488 A; 331 C; 482 G; 498 T; 0 U; 0 Other;

  Query Match             19.9%;  Score 1487.1;  DB 1;  Length 1799;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches 1613;  Conservative   27;  Mismatches  116;  Indels   45;  Gaps    5;

Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||||| ||||| |||||||:|||||||||||||||||
Db       1799 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 1741

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||| ||||||
Db       1740 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 1681

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
               ||||| ||||||||||:|||:|||||||||||| |||||||| || |||| |||||||||
Db       1680 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 1621

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
             ||||||||||  |||||||| |  :||| |||| | ||||||||||||||||||||||||||
Db       1620 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 1561

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
             |||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 1501

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
             |||| |||||||||||||||||||||||||| ||||||||| |||||||||| ||||||
Db       1500 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 1441

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
             |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1440 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 1381

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAG 6022
             ||||||| ||||||| ||||||||| |||:|||||||| ||||| ||||||||||||||||
Db       1380 ATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCTACAAAG 1321

Qy       6023 GACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTAT 6082
             |||||||||||| |||||||||||||||||||| |||||||||| ||||||||||||||||
Db       1320 GACTAGATCTCTTAAAACTACATGAAACCCTCCATACCCATACTTGCCTGGTAAGCCTAT 1261

Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGA 6142
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1260 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAAACCCTACTAACTGTTGGT 1201

Qy       6143 TATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACA 6202
             | ||||||||||| | ||||:|||||   ||||||||||||||||||||||||||||||||
Db       1200 TGTGCCTCCCCCTGTATTTCAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACA 1141

Qy       6203 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1140 ACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGG 1081

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
             ||||||||||||||||||||||||||||||||||||||||||||||||  | |||||||||
Db       1080 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCA 1021

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAA 961

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        960 TATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 901

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
             |||||||||||||||||||||||||| |||:|||| |||||||||||||||||||||| ||
Db        900 GCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATT 842
```

```
Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
Db        841 ATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAG 782

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
Db        781 GAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACT 722

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
Db        721 ACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCT 662

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
Db        661 TGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTAGACT 602

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGG---------- 6782
Db        601 TGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATG 542

Qy       6783 ----------GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAG 6832
Db        541 TTATTTTAGCGGAAGAATGTTGTTATTATGTTAATCAATCCTGAATTGTCACAGAGAAAG 482

Qy       6833 TTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC-GAAACACTGGACCC 6891
Db        481 TTGAAGAAATTCGAGATTGAATACAACGTAGAACAGAGGAGCTTCAAAAACACCAGACCC 422

Qy       6892 TGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCT 6951
Db        421 TGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGATCTCTAGCAGCT 362

Qy       6952 ATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCT 7011
Db        361 CTAATATTGATACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCT 302

Qy       7012 TCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCC 7059
Db        301 TCCAGAATCAAAGTTGTAAAGCTACAAATCGTTCTTCAAATGGAACCCCAGATGAAGTCC 242

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
Db        241 ATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTACTAGCCCATGCTCCAATTGTAAT 182

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
Db        181 GATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAAT 122

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
Db        121 TCCGCAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCCAACAGCACTTGGGTTTTCC 62

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
Db          61 TGTTGAGTGGGGGGACTGAGAGACAGGATTAGCTGGATTTCCTAGGCCGACTAAGAATCC 2

Qy       7300 Y 7300
              :
Db          1 C 1
```

```
RESULT 46
AAX25659
ID   AAX25659 standard; cDNA to mRNA; 1948 BP.
XX
AC   AAX25659;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.24.4.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 56-58; 106pp; French.
XX
CC   This sequence represents clone cl.24.4 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
Db         491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
Db         551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
Db         611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
Db         671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
Db         730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
Db         790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
Db         850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
Db         910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
```

```
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
              ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
              ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
              ||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948


RESULT 47
AAA59209
ID   AAA59209 standard; DNA; 1948 BP.
XX
AC   AAA59209;
XX
DT   07-NOV-2000  (first entry)
XX
DE   5' non coding, 3' pol gene, and 5' env gene of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 45; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;

  Query Match           19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
              ||||||||||||||||:||:||||:|||:||||||||||||||||||||||||||||||
Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
```

```
Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
              ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
              ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
              ||||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948
```

```
RESULT 48
AAA96625
ID   AAA96625 standard; DNA; 1629 BP.
XX
AC   AAA96625;
XX
DT   08-FEB-2001  (first entry)
XX
DE   DNA encoding an envelope (env) protein of MSRV-1.
XX
KW   Envelope protein; MSRV-1; superantigen; autoimmune disease; Vbeta16;
KW   Vbeta17; multiple sclerosis; vaccine; ss.
XX
OS   Human endogenous retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1629
FT                   /*tag=  a
FT                   /product= "envelope protein"
XX
PN   WO200057185-A1.
XX
PD   28-SEP-2000.
XX
PF   20-MAR-2000; 2000WO-FR000691.
XX
PR   19-MAR-1999;   99FR-00003622.
PR   28-OCT-1999;   99FR-00013755.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Lafont M;
XX
DR   WPI; 2000-638279/61.
DR   P-PSDB; AAB19069.
XX
PT   Detecting superantigen activity, useful for identifying agents for
PT   treatment or prevention of autoimmune disease, from expansion or loss of
PT   particular lymphocyte Vbeta determinants.
XX
PS   Claim 25; Page 123-124; 134pp; French.
XX
CC   The present sequence encodes an envelope protein of MSRV-1. The envelope
CC   protein expressed by the endogenous human retrovirus MRSV-1 has
CC   superantigen activity associated with autoimmune disease. The protien can
CC   be detected using the method of the invention. The specification
CC   describes a process for detecting activity of a superantigen in a
CC   biological sample. The process comprises identifying large scale
CC   expansion or loss of lymphocytes that carry at least one of the Vbeta16
CC   and/or Vbeta17 determinants. The method is used to screen for agents that
CC   inhibit the superantigen, especially those associated with MRSV-1 which
CC   is implicated in autoimmune disease, particularly multiple sclerosis.
CC   These agents are potentially useful for treatment or prevention (e.g. as
CC   vaccines) of autoimmune diseases
XX
SQ   Sequence 1629 BP; 456 A; 454 C; 287 G; 432 T; 0 U; 0 Other;

  Query Match             19.1%;  Score 1431;  DB 1;  Length 1629;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches 1492;  Conservative   25;  Mismatches  100;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:|||||||||||||| |||||||||:||:||||:| |:|||| :|||| |||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||| ||||||  |||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              | || || ||| |||||||||||||||||||| ||||||||| | |:|||||||| | ||
```

```
Db       121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
Qy      5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db       181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy      5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db       241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy      5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db       301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy      5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db       361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy      6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db       421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy      6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db       481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy      6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db       541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy      6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db       601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy      6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db       661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy      6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db       721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy      6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db       781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy      6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db       841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy      6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db       901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy      6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db       961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy      6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db      1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy      6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db      1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy      6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db      1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy      6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db      1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy      6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db      1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy      6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db      1321 CTCAGCCAATGGATGCCCTGGACTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTT 1380

Qy      6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db      1381 TTACTCCTCTTTGGACCCTGTATCTTCAACTTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy      7021 GAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
Db      1441 GAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAACCCCAGATGCAGTCCATGACTAAA 1500

Qy      7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
Db      1501 ATCTACCGTGGACCCCTGGACCGGCCTGCTAGACTATGCTCTGATGTTAATGACATTGAA 1560

Qy      7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
Db      1561 GTCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTACACTCCAATTCAGTAGGA 1620

Qy      7189 AGCAGTTAG 7197
Db      1621 AGCAGTTAG 1629
```

RESULT 49
AAS73701
ID  AAS73701 standard; cDNA; 1715 BP.
XX
AC  AAS73701;
XX
DT  13-FEB-2002  (first entry)
XX
DE  DNA encoding novel human diagnostic protein #9505.
XX
KW  Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW  food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS  Homo sapiens.
XX
PN  WO200175067-A2.
XX
PD  11-OCT-2001.
XX
PF  30-MAR-2001; 2001WO-US008631.
XX
PR  31-MAR-2000; 2000US-00540217.
PR  23-AUG-2000; 2000US-00649167.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Drmanac RT,  Liu C,  Tang YT;
XX
DR  WPI; 2001-639362/73.
DR  P-PSDB; ABG09514.
XX
PT  New isolated polynucleotide and encoded polypeptides, useful in
PT  diagnostics, forensics, gene mapping, identification of mutations
PT  responsible for genetic disorders or other traits and to assess
PT  biodiversity.
XX
PS  Claim 1; SEQ ID NO 9505; 103pp; English.
XX
CC  The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC  sequences. (I) is useful as hybridisation probes, polymerase chain
CC  reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC  and in recombinant production of (II). The polynucleotides are also used
CC  in diagnostics as expressed sequence tags for identifying expressed

```
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1715 BP; 512 A; 435 C; 371 G; 397 T; 0 U; 0 Other;

  Query Match             18.5%;  Score 1383.7;  DB 1;  Length 1715;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches 1551;  Conservative   29;  Mismatches   97;  Indels   39;  Gaps   11;

Qy         90 CTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCG 149
Db          1 CTATTTCACTCTATTAAATCTTGCAACTGCATTCTTCTGGTCCGTGTTTGTTACAGCTCG 60

Qy        150 AGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTT----------------GCCACCA 192
Db         61 AGCTGAGCTTTCACTTGCTGCCCACCACTGGTGTTTGCCGCTGTCACAGACCCGCCGCTG 120

Qy        193 CCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCT 252
Db        121 TCGCAGACCCGCTGCTGACTGCCATCCGTCTGGATCCGGCAGGGTGTCCACTGTGCTCCT 180

Qy        253 GATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGC 312
Db        181 GATCCAATGAGGCACCCATTGCCGCTCCCGATCCGGCTAAAGGCTTGCCATTGTTCCTGC 240

Qy        313 ACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGT 372
Db        241 ACAGCTAGGTGCCCGGGTTTGTCCTAATCAAGCTGAACACTACTTGCTGGGTTCCACGGT 300

Qy        373 TCTCTTCTGTGACCCACGGCTTCTAAT---------AKAACTATAACACTTACCACATGG 423
Db        301 TCTCTTCCGTGACCCACGGCTTCTAATAGAGCTAACAGAGCTATAACACTCACCGCACGG 360

Qy        424 CCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGAR 483
Db        361 CCCAAGATTCCATTCCTTCGAATCCGTGAGGCCAAGAATCCCCAGGTCAGAGAATACGAA 420

Qy        484 GCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTG 543
Db        421 GCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTG 480

Qy        544 GGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAA 603
Db        481 GGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAA 540

Qy        604 GTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGG 663
Db        541 GTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAATTGG 600

Qy        664 GAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCC 723
Db        601 GAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCC 660

Qy        724 TGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAG 783
Db        661 TGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAG 720

Qy        784 AAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
Db        721 AAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 780

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
Db        781 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 840

Qy        904 TCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCC 963
Db        841 TCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGGTCC 900

Qy        964 AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTAT 1023
Db        901 AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTAT 960

Qy       1024 GACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTT 1083
Db        961 GACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCTTTT 1019

Qy       1084 YYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGAT 1143
Db       1020 TCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCTGAT 1079

Qy       1144 GGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA 1203
Db       1080 GGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA 1139

Qy       1204 TATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGC 1263
Db       1140 TATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGC 1199

Qy       1264 CTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAA 1323
Db       1200 CCGAGAGCTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-A 1256

Qy       1324 GGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGG 1383
Db       1257 GGAAAGAGAACGATTCCCTACAGGCCAGCAGGCAGTTCCCAGTGCAGACCCTCATT---G 1313

Qy       1384 ACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTAS 1443
Db       1314 AGACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTACTAACTTGCGTGCTGG 1372

Qy       1444 AAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACA 1503
Db       1373 AAGGACTAAGGAAAACCAGGAAG-AAGGCTGTGAATTTTTCAGTGATGTCCACTGTAACA 1431

Qy       1504 CAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAA 1563
Db       1432 CA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCACTGAGGAA 1490

Qy       1564 GCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTAT 1623
Db       1491 GCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGATTAAGTTTAT 1550

Qy       1624 CACTCAGTCAGCTGCAGACATTAGAAA--AAACTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
Db       1551 CACTCGGTCAGCTGCAGACATTAGAAAACAAACTTCAAAAGTCCACCTTAGGCCTGGAGC 1610

Qy       1682 AAAACTT-AGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAG 1740
Db       1611 AAAACTTAAGACACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAG 1670

Qy       1741 GAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA 1776
Db       1671 GAGCAGTCGGAATGGGACAAACGAGATTAAAAAAAA 1706


RESULT 50
AAX25657
ID   AAX25657 standard; cDNA to mRNA; 1422 BP.
XX
AC   AAX25657;
XX
```

```
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.7A16.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 52-54; 106pp; French.
XX
CC   This sequence represents clone cl.7A16 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 1422 BP; 448 A; 312 C; 316 G; 344 T; 0 U; 2 Other;

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||:||||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||||||:|||||:||||||||||||:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||| :|:|:|||||:||||||||||||||||||||||||||||:||||:|||||
Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             |||||||-|||||:|||||||||||||||||:::|||||||||||||||||||||||||||
Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||
Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             ||||||||||:||||:|||||||||||:|||||:|||||:||||||||||||||||||||
Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||||||||||:|||||||||:|:||||||:::|||||||||||||||||||||||
Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             ||||||||||||||||||:|||||||||||:|||||||:|||||||||||||||:||:||
Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |::|||||||||||||||||:||:||||||:|||||||||||||||:|||:|||||:||||
Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
             ||||||||::||||::||||||:||||||||:||||||||||||||||||:|:|||||
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||:|||||||||||:|||||||||||||||||||:|||||
Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:|||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
                ||||||:||:|||||||||||||:|||||||||||||||||||||||||||||||||||
Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:|||||||||||:|||||||||||||||||||||||||||||||||||
Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
             |||||||||||||||| |||||:|||||||||||||||||||||||||||||:||||
Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
             |||| ||||||||||||||||||||||||||:||||:|||||||||||||:|||:|:|
Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
             ||||||||||||||||||||:|||||||:|||||||||||||||||||:|:||||||:|
Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
             ||||||||||||||||||::||||||||:|:|||||||||||||||||:|||:|
Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227
             ||||||||||:|||||||||||||||||||:|||||||:||||:|||||||||||||||:
Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287
             ||||||||:|||||||||||:|||||||||| |||||||||:|||||||||||:|||:
Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337
             |||:|||||||||:||||||||||:|||||||||||||||||||||||||
Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 51
AAA59207
ID   AAA59207 standard; DNA; 1422 BP.
XX
```

```
AC   AAA59207;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Fragment 3' of the pol gene of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 44; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 1422 BP; 448 A; 312 C; 316 G; 344 T; 0 U; 2 Other;

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
Db        1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
Db        1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
Db        1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
Db        1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227
Db        1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287
Db        1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337
Db        1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422
```

```
RESULT 52
AAV43217
ID   AAV43217 standard; cDNA; 1481 BP.
XX
AC   AAV43217;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus fragment 5.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1479
FT                   /*tag=  a
FT                   /product= "Encodes protein AAW71068"
FT                   /transl_except= (pos:115-117, appears to code for a stop
FT                   codon)
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71068.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 184-185; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) genomic fragment used in the method of the invention.
CC   The invention provides complete or partial genomic sequences of the MSRV-
CC   1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC   genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             |||||||||||||| |||||| ||||||||||||||:|||:|||||||||||| |||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             | || || ||| |||||||||||||||||||| ||||||||| | |:||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             |||| |||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             |||| |||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             |||||||||||||| |||||||||||||||||||||||| |||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             |||||||||||||||||||||| |||||||||||| ||||||:|||||| ||||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||| | |||||||||| |||||||||||||||||||||||| |||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             |||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             |||||||||||||||||| | ||||| || ||||||||| ||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||  ||||||||||||||||  |||||||||||||||||||| |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||||| |||||| |||||||||| |
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:|||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||| |||| ||||||||||||||:|||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:|||||||:||||||| |||||| | || ||||||||||
```

```
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| |||||||||||||||||||||||||||||| ||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              |||||||||||||||||||||||||:||||||||||||| |||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
              ||||||||:|| ||||| |||| |||  | ||| |  || |
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481


RESULT 53
AAX29703
ID   AAX29703 standard; DNA; 1481 BP.
XX
AC   AAX29703;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Clone C15 from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis; ss.
XX
OS   Viruses.
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   P-PSDB; AAW99553.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 1; Page 37-38; 83pp; French.
XX
CC   This sequence represents clone C15 from a novel multiple sclerosis
CC   related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC   prophylactic or therapeutic compositions to inhibit expression of a
CC   multiple sclerosis related virus and/or virus associated with rheumatoid
CC   polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC   17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| |||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||| ||||||  |||||||||||||||:|||:|||||||||||| |||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              | || || ||| |||||||||||||||||||| ||||||||| | |:||||||| | ||

Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||| |||||||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||| |||||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||| |||||||||||||||||||||| |||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||||||| |||||||||||| ||||||:|||||  ||||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||| | ||||||||| |||||||||||||||||||||||||| |||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              |||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              |||||||||||||||| | ||||| || ||||||||| ||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              |||||||||||||| ||||||||||||||| ||||||||||||||||||||| |||||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||||| |
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              |:|||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
              || |||||| |||||||||||||||||||||| |||| |||||||||||||||:|||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260
```

```
Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
```

RESULT 54
ADG14846
ID ADG14846 standard; cDNA; 1481 BP.
XX
AC ADG14846;
XX
DT 26-FEB-2004 (first entry)
XX
DE MSRV associated cDNA #9.
XX
KW ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003198647-A1.
XX
PD 23-OCT-2003.
XX
PF 03-APR-2002; 2002US-00114104.
XX
PR 26-NOV-1996; 96US-00756429.
PR 26-NOV-1997; 97US-00979847.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2004-032461/03.
XX
PT New isolated nucleic acid and their fragments having the pol gene of a
PT retrovirus, useful for diagnosing, preventing and/or treating multiple
PT sclerosis and/or rheumatoid arthritis.
XX
PS Disclosure; SEQ ID NO 105; 193pp; English.
XX
CC The invention relates to an isolated nucleic acid which comprises the pol
CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC arthritis. The methods and compositions of the present invention are
CC useful for diagnosing, preventing and/or treating multiple sclerosis
CC and/or rheumatoid arthritis. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;

Query Match 17.7%; Score 1324.2; DB 1; Length 1481;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;

```
Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
```

```
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200
```

```
Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||||||| |||| ||||||||||||||:|||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:||||||:||||||| |||||| | || ||||||||||
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| |||||||||||||||||||||||||||||| ||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
               ||||||||||||||||||||||||:||||||||||||| ||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
             ||||||||:|| ||||| |||| | ||| | || |
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
```

```
RESULT 55
ADB84502
ID   ADB84502 standard; DNA; 1600 BP.
XX
AC   ADB84502;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol gene consensus sequence #3.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 52A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 1600 BP; 458 A; 383 C; 322 G; 390 T; 0 U; 47 Other;
```

```
  Query Match             17.6%;  Score 1321.4;  DB 1;  Length 1600;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches 1403;  Conservative  111;  Mismatches   86;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||||  | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| |||||||   |||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |         |||   | | | |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||||| |||  |||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:||||||:|:|||||| :|||||||||||||||:|:::
Db        650 GYCARTYYTCATACCTGGACAYTCTTGTYCYTCRGTACRTGGATGATTTACTTTTAGYCR 709
```

```
Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||||||||||||||||||||::|||||:|:||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGMACTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:|:||||::|||||||||||||:||:|||:||||||||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCRGCTCTGCTCACAGSAGRTTARATACTTAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||||||:::|||||||||||||||:::|||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGRRCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||||||||:|||||||||||||||||||::|::|||||||||:|||:||||:||||
Db        890 STTATCCTCATCCCAAAACCCTAAAGCAACTAASAGSGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             |||:||||::|||| |||||||:  :|::::|||||||  :|||||||||||::|||||:|
Db        950 GCCRAAWAYRGATT-CCCAGGTACASCMMRRTAGCCAGACCATTAWATACACKAATTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             ||||||||:|||||||:||||:|||||:|||||||||:|:||||||:::|||||||||||
Db       1009 GAAACTCARAAAGCCARTACCYATTTAGTAAGATGGACAYCTGAAGCAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  ||||||||:|||||||| ||:||||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
               |||||:|||||||||||||::||| ::|||||||:|||||||||:|||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:||||||||||:||||| ||| ||||||||||:: |||||||||||||||||||
Db       1246 GGTTGGCCTCATTGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:||||||||| |||||||||||||||||||:|||||::||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| |||||||||:  |||||||| |:||||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
               ||||||||||||||:||||:|||||:||||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 56
ADG14948
ID   ADG14948 standard; cDNA; 1600 BP.
XX
AC   ADG14948;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #30.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 207; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1600 BP; 458 A; 383 C; 322 G; 389 T; 0 U; 48 Other;

  Query Match           17.6%;  Score 1320.4;  DB 1;  Length 1600;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches 1402;  Conservative  111;  Mismatches   87;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||||  | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| ||||||| |||||||||:|||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |         |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||:|||||||| |||||||||||||||| ||||||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529
```

```
Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             ||||||||||||||||||||||||| |||  |||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:||||||:|:|||||| :|||||||||||||||:|:::
Db        650 GYCARTYYTCATACCTGGACAYTCTTGTYCYTCRGTACRTGGATGATTTACTTTTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||||||||||||||||||||||::|||||:|:|||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGMACTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:||||||||||||:|:|:||||::|||||||||||||:||:|||:||||||||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCRGCTCTGCTCACAGSAGRTTARATACTTAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||||||||:::|||||||||||||||:::|||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGRRCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||||||||:|||||||||||||||||||||||::|::|||||||||:|||:||||:||||
Db        890 STTATCCTCATCCCAAAACCCTAAAGCAACTAASAGSGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             |||:||||::|||| |||||||:  :|::::|||||||  :|||||||||||::|||||:|
Db        950 GCCRAAWAYRGATT-CCCAGGTACASCMMRRTAGCCAGACCATTAWATACACKAATTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             ||||||||:|||||||:||||:|||||:|||||||||:|:||||||:::|||||||||||
Db       1009 GAAACTCARAAAGCCARTACCYATTTAGTAAGATGGACAYCTGAAGCAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  ||||||||:|||||||| ||:|||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:||||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
               |||||:|||||||||||||::||| ::||||||:||||||||:|||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:|||| |||||:||||| ||| |||||||||||:: ||||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:|||||||| ||||||||||||||||||:|||||::||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| ||||||||||: |||||||| |:||||||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||||||||||:||||:|||||:||||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600
```

```
RESULT 57
ADG14947
ID   ADG14947 standard; cDNA; 1600 BP.
XX
AC   ADG14947;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #29.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 206; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1600 BP; 453 A; 377 C; 321 G; 381 T; 0 U; 68 Other;

  Query Match           17.5%;  Score 1309.8;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db         120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||||  | |||||||||||||||||| ||||||| || |
Db         179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| |||||||   ||||||||||||||||| |||||||||||||| |||
Db         239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||| |  ||||||| || | |         |||   |  | |  |||||||
Db         299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351
```

```
Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             |||||||:|||||||||:|||||||||| |||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| ||| ||||||:| ||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:||||||:|:||:|||:::||||||:|||:||:|:|:::
Db        650 GYCARTYCTCATACCTGGACAYTCTTGTYCYTCAGTAYRKGGATGAYTTAMTTWTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||||:|| |||||||||||||:|||||||:|:|||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGSCAMCAAGCCACCCAAGHRCTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:::||||::|||||||||||::|||:|||:|||||  ||
Db        770 GTGGCTACAAGGTTTCCAAACMAARGGCTCRSCTCTGCTCACASSAGRTTARATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:||||||||||||:|:|||:|||||||||:|||||:::|||||||:|||||||||
Db        830 GGCTAAAATTATCCAAAGKCRCCAGGGCCCTCAGWGAGGAACGTATCCAGCSTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||| ||||:||:||||:|||||||||||||:::||||||||||:||::||:|||||
Db        890 MTTATCCMCATCCCAWAACCMTAAAGCAACTAAGARGGTTCCTTGGCATAWCAGSYTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||||::|||| |||:|:|:  ::::|||||||||::|||||::||||:||::||:|
Db        950 GCCGAAWAYRGATT-CCCVGRTACASYSMAATAGCCAGRCCATTATRTACAYTADYTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||||:||:||:|||||||||:|:||||: :::|||||||||||
Db       1009 GAAACTCAGAAAGCCAATACCYATWTAGTAAGATGGACACCTGARACAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  ||||||||:|||||||| ||:|||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
               |||||:|||||||||||::||| ::|||||:||||||||:|||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:|||| |||||:||||| ||| ||||||||||:: ||||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:||||||||| ||||||||||||||||||:|||||::||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| ||||||||||: ||||||||| |:|||||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||||:||:|||||||||||||||:||||| ||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 58
ADB84501
ID   ADB84501 standard; DNA; 1600 BP.
XX
AC   ADB84501;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol gene consensus sequence #2.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 51A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
```

```
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 1600 BP; 453 A; 377 C; 320 G; 383 T; 0 U; 67 Other;

  Query Match             17.5%;  Score 1309.2;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||| | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| |||||||   |||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |        |||    |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  ||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||||| |||  |||||||:| ||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:||||||:|:||:|||:::||||||:|||:||:|:|:::
Db        650 GYCARTYCTCATACCTGGACAYTCTTGTYCYTCAGTAYRKGGATGAYTTAMTTWTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||||:|| |||||||||||||:|||||||:|:|||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGSCAMCAAGCCACCCAAGHRCTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:::||||::|||||||||||::||:|||:|||||  ||
Db        770 GTGGCTACAAGGTTTCCAAACMAARGGCTCRSCTCTGCTCACASSAGRTTARATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:||||||||||||:|:|||:|||||||||:||||||:::||||||||:||||||||

Db        830 GGCTAAAATTATCCAAAGKCRCCAGGGCCCTCAGWGAGGAACGTATCCAGCSTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||| ||||:||:||||:|||||||||||||:::|||||||||:||::||:|||||
Db        890 MTTATCCMCATCCCAWAACCMTAAAGCAACTAAGARGGTTCCTTGGCATAWCAGSYTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||||:: ||| |||:|:|:  ::::|||||||||::|||||::||||:||::||:|
Db        950 GCCGAAWAYRTATT-CCCVGRTACASYSMAATAGCCAGRCCATTATRTACAYTADYTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||||:||:||:|||||||||:|:||||: :::|||||||||||
Db       1009 GAAACTCAGAAAGCCAATACCYATWTAGTAAGATGGACACCTGARACAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             ||||||  ||  | | |||  ||||||||:|||||||| ||:|||||||:|||||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:|||||||:|||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
              |||||:||||||||||::||| ::||||||:||||||||:||||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:||||||||||:||||| ||| |||||||||||:: ||||||||||||||||
Db       1246 GGTTGGCCTCATTGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:|||||||| |||||||||||||||||||:|||||::|||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| |||||||||||: |||||||| |:|||||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

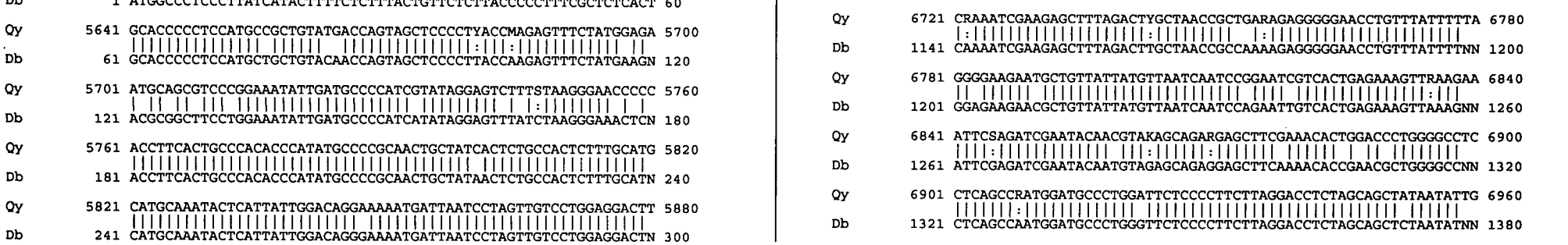
Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||||:|||||  ||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 59
ADB84400
ID   ADB84400 standard; DNA; 1481 BP.
XX
AC   ADB84400;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #19.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
```

XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Claim 31; Page 77-78; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.
XX
SQ Sequence 1481 BP; 405 A; 403 C; 257 G; 385 T; 0 U; 31 Other;

Query Match 17.3%; Score 1296.4; DB 1; Length 1481;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1336; Conservative 24; Mismatches 121; Indels 0; Gaps 0;



```
Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAANN 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481


RESULT 60
ADG14946
ID   ADG14946 standard; cDNA; 1597 BP.
XX
AC   ADG14946;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #28.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 205; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1597 BP; 468 A; 393 C; 331 G; 395 T; 0 U; 10 Other;

  Query Match             17.2%;  Score 1291.9;  DB 1;  Length 1597;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches 1405;  Conservative   88;  Mismatches  104;  Indels   19;  Gaps   10;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
Db         239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
Db         299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
Db         352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
Db         410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
Db         470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
Db         530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
Db         590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
Db         650 GTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCG 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
Db         710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
Db         770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
Db         830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
Db         890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
Db         950 GCCGAAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
Db         1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
Db         1066 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1123

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
Db         1124 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1182

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
Db         1183 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1242

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
Db         1243 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1302
```

```
Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             ||||||||||||||:|||||||| ||||||||||||||||||||:|||||::||||||||
Db       1303 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1362

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| ||||||||||: ||||||||| |:|||||||||||||
Db       1363 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1422

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||| || |:| ||||||||
Db       1423 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1482

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||
Db       1483 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1542

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||||||||||:||||:|||||:|||||||||||:|||| ||||||||||||
Db       1543 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1597
```

RESULT 61
ADB84500
ID ADB84500 standard; DNA; 1597 BP.
XX
AC ADB84500;
XX
DT 04-DEC-2003 (first entry)
XX
DE MSRV-1 pol gene.
XX
KW MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003039664-A1.
XX
PD 27-FEB-2003.
XX
PF 26-NOV-1997; 97US-00979847.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Example 19; Fig 51A; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC The SEQ ID numbers for the sequences as displayed in the main body of the
CC patent do not match the SEQ ID numbers in the sequence listing.
CC Consequently those sequences mentioned in the claims may not be the
CC sequences the authors intended to claim.
XX
SQ Sequence 1597 BP; 469 A; 393 C; 331 G; 395 T; 0 U; 9 Other;

Query Match 17.2%; Score 1288.1; DB 1; Length 1597;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1403; Conservative 88; Mismatches 106; Indels 19; Gaps 10;

```
Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| ||||||||||||||||||| |||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||||| | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| |||||||  |||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  | ||||||| || | |         |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| |||  ||||||:| |||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |:||:|::|||||||||||||||||||||||:|||  |||||||||||||||:| :|
Db        650 GTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCG 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
```

```
Db        710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135

Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195

Db        830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255

Db        890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315

Db        950 GCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375

Db       1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435

Db       1066 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1123

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495

Db       1124 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1182

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555

Db       1183 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1242

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615

Db       1243 GGTTGGCCTCATNGTTATTGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1302

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675

Db       1303 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1362

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735

Db       1363 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1422

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795

Db       1423 CTTGAAGAGCCAGTGCTGAGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1482

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855

Db       1483 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1542

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911

Db       1543 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1597


RESULT 62
AAS76459
ID   AAS76459 standard; cDNA; 1660 BP.
XX
AC   AAS76459;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12263.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12272.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 12263; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1660 BP; 511 A; 396 C; 351 G; 402 T; 0 U; 0 Other;

  Query Match           17.2%;  Score 1285.7;  DB 1;  Length 1660;
  Best Local Similarity   77.2%;  Pred. No. 0;
  Matches 1470;  Conservative   92;  Mismatches   98;  Indels  245;  Gaps    6;

Qy       2522 GCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAA 2581

Db          1 GCAGTACCCCCTTAGACCCAAGGCCCAACAAGGACTCCAAAAAACTGTTAAGGACCTAAA 60

Qy       2582 AGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGG 2641

Db         61 AGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAATACTCC-------AACTGTAG 113

Qy       2642 AGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGA 2701

Db        114 GAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAAATCTCAGGATTATCAAT-GA 171

Qy       2702 GGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACC 2761

Db        172 GACTGTTGTTCCTCTATACCCAGCTGAACCTAACCCTTATACTCTGCTTTCCCAAATACC 231

Qy       2762 AGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGT 2821

Db        232 AGAGGAAGCAGAGTGGTATATAGTCCTGGACCTTAAGGATGCCTTCTTCTGCATCCCTGT 291
```

```
Qy      2822 ACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACT 2881
             |||||||||  ||  ||||||  ||||||||||||||||||  ||||||||  ||:|  ||||||||
Db       292 ACATCCTGATTCCCAATTCCTGTTTGCCTTTGAAGATCCTTCAAACTCAACGTCTCAACT  351

Qy      2882 CACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATT 2941
             ||||||||||:||||||  ||||||||||||||||||:|||||||||||||||||||||||
Db       352 CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATT  411

Qy      2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             |||||||||  |||||:||:|::|||||||||||||||||||||||| :|||:||||||||
Db       412 AGCCCAAGATTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGA  471

Qy      3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
             |||||||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:||
Db       472 TTTACTTTTAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAA  531

Qy      3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
             |||||||  |:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||
Db       532 TTTCCTCACCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCA  591

Qy      3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
             ||||||||||||||:||||:|||||||||||||   |:|||||||||||||||:::||
Db       592 GGTTAAATACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACAT  651

Qy      3182 CCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGG 3241
             |||||||||||||||||||||||||||:|||||||||||||||||||||||:::||||||||
Db       652 CCAGCCTATACTGGCTTATCCTCTTCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGG  711

Qy      3242 CRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAW 3301
             |:|||:|||:|||| |||||: |||||| |||||||:  ||:|||||| ||| :|||||:
Db       712 CATAACAGGTTTCTACCGAATGTGGATT-CCCAGGTACAGCAAAATAGGCAGATCATTAT  770

Qy      3302 ATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAG 3361
             |||||:|||||||  ||||||||||||||||||||||||||
Db       771 ATACACTAATTAAAGAAACTCAGAAAGCCAATACCCATTTA------------------- 811

Qy      3362 YMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTT 3421
Db       812 ------------------------------------------------------------ 811

Qy      3422 GCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAG 3481
Db       812 ------------------------------------------------------------ 811

Qy      3482 TCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTG 3541
Db       812 ------------------------------------------------------------ 811

Qy      3542 ATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAG 3601
                                                 ||||||||||||||||||||::|||
Db       812 ----------------------------------GTAGTGGTGGCAGTAGCAGTCTTAG  836

Qy      3602 TATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKG 3661
             | |||||||  ||||||||  ||||||||:||||||||||||||||||||||||||:|
Db       837 TGTCTGAAGTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATG  896

Qy      3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
             ||||:: |||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       897 TGAATGCCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAAT  956

Qy      3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
             :|||||||||||||||||:|||||||||||:|||| |||||||||| ||||||||||||
Db       957 ATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACC 1016

Qy      3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTT 3841
             |||:| ||||||||||||||||||||||||||||:||:|||||||| ||||||:|||||
Db      1017 CAGTCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTG 1076

Qy      3842 CTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTC 3901
             |||||||||||||| |||||||||||  ||:||||:|||||:|||||||||||: ||| ||
Db      1077 CTCAAACCTATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TC 1135

Qy      3902 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1136 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCA 1195

Qy      3962 GTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAG 4021
             ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||:||
Db      1196 GTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAG 1255

Qy      4022 CTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAAT 4081
             ||:|||||||||||||:|||:| ||| |||||||| ||||||||:||||||||:||||
Db      1256 CTGGCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAAT 1315

Qy      4082 ATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARA 4141
             ||||||||||||||||:|:| ||||:|||||||| ||||||||||||::|||||||:|:|
Db      1316 ATATATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGA 1375

Qy      4142 GAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARA 4201
             |||||||||||||||||||||:|||:|||||||||||:|||||||||||||||||||:|
Db      1376 GAAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAA 1435

Qy      4202 TTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCAN 4261
             ||||||:||||:||||||||||||:|||||||:||| |||||||||:|||||||||||
Db      1436 TTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAG 1495

Qy      4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
             |||||||||:|||||||||||:|||:|| :  |||||  :|||||||| :|||||||||
Db      1496 AAAGGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCC 1555

Qy      4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
             ||||||  ||||||||||||||||||||||||||||||    |   ||| ||||||  ||||||||||
Db      1556 ACAAGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCC 1615

Qy      4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
             ||||||||  |||||||||||||||||||||||||||||||||| ||||| |
Db      1616 CTCCGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAG 1660
```

RESULT 63
AAS73608
ID AAS73608 standard; cDNA; 1660 BP.
XX
AC AAS73608;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9412.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG09421.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 9412; 103pp; English.

```
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1660 BP; 511 A; 396 C; 351 G; 402 T; 0 U; 0 Other;

  Query Match             17.2%;  Score 1285.7;  DB 1;  Length 1660;
  Best Local Similarity   77.2%;  Pred. No. 0;
  Matches 1470;  Conservative   92;  Mismatches   98;  Indels  245;  Gaps    6;

Qy       2522 GCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAA 2581
             |||||||||||| |||||||||||||||||||| ||||||| | |||||||||| ||||
Db          1 GCAGTACCCCCTTAGACCCAAGGCCCAACAAGGACTCCAAAAAACTGTTAAGGACCTAAA 60

Qy       2582 AGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGG 2641
             |||||||||| ||||||||||||||   |  ||||||| || | |         | ||  |
Db         61 AGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAATACTCC-------AACTGTAG 113

Qy       2642 AGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGA 2701
               | ||||||||||  |||||||||| |||||||||| ||||||||||||||||||  ||
Db        114 GAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAAATCTCAGGATTATCAAT-GA 171

Qy       2702 GGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACC 2761
             | | ||||| | | |||||||||||| ||||| |||||||||| ||::|||||||||||||
Db        172 GACTGTTGTTCCTCTATACCCAGCTGAACCTAACCCTTATACTCTGCTTTCCCAAATACC 231

Qy       2762 AGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGT 2821
             |||||||||||||||| || |:|||||||||||:|||||||||||||||||||||||||||
Db        232 AGAGGAAGCAGAGTGGTATATAGTCCTGGACCTTAAGGATGCCTTCTTCTGCATCCCTGT 291

Qy       2822 ACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACT 2881
             ||||||||| || |||||| |||||||||||||||| |||||||| ||:| ||||||||
Db        292 ACATCCTGATTCCCAATTCCTGTTTGCCTTTGAAGATCCTTCAAACTCAACGTCTCAACT 351

Qy       2882 CACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATT 2941
             ||||||||||:|||||| |||||||||||||||||:|||||||||||||||||||||||||
Db        352 CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATT 411

Qy       2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             ||||||||| |||||:||:|::|||||||||||||||||||||||| :|||:|||||||||
Db        412 AGCCCAAGATTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGA 471

Qy       3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
             |||||||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:||
Db        472 TTTACTTTTAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAA 531

Qy       3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
             ||||||| |:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||
Db        532 TTTCCTCACCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCA 591

Qy       3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
             |||||||||||||||:||||:|||||||||||||   |:|||||||||||||||:::||
Db        592 GGTTAAATACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACAT 651

Qy       3182 CCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGG 3241
             |||||||||||||||||||||||||||||| ||:|||||||||||||||||||||:|::|||||||
Db         652 CCAGCCTATACTGGCTTATCCTCTTCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGG 711

Qy       3242 CRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAW 3301
             |:|||:|||:|||| |||||: |||||| |||||||:  ||:|||||| ||| :|||||:
Db         712 CATAACAGGTTTCTACCGAATGTGGATT-CCCAGGTACAGCAAAATAGGCAGATCATTAT 770

Qy       3302 ATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAG 3361
             |||||:|||||| |||||||||||||||||||||||||||
Db         771 ATACACTAATTAAAGAAACTCAGAAAGCCAATACCCATTTA---------------------- 811

Qy       3362 YMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTT 3421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         812 ------------------------------------------------------------ 811

Qy       3422 GCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAG 3481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         812 ------------------------------------------------------------ 811

Qy       3482 TCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTG 3541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         812 ------------------------------------------------------------ 811

Qy       3542 ATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAG 3601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||::|||
Db         812 ----------------------------------GTAGTGGTGGCAGTAGCAGTCTTAG 836

Qy       3602 TATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKG 3661
             | |||||||  |||||||| |||||||:|||||||||||||||||||||||||||||:|
Db         837 TGTCTGAAGTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATG 896

Qy       3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
             ||||:: |||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db         897 TGAATGCCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAAT 956

Qy       3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
             :|||||||||||||||||:|||||||||||:|||| |||||||| |||||||||||||||
Db         957 ATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACC 1016

Qy       3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTT 3841
             |||:| |||||||||||||||||||||||||||||||:||:|||||||| ||||||:|||||
Db        1017 CAGTCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTG 1076

Qy       3842 CTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTC 3901
             |||||||||||||| |||||||||| ||:||||:|||||:|||||||||||: ||| ||
Db        1077 CTCAAACCTATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TC 1135

Qy       3902 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||
Db        1136 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCA 1195

Qy       3962 GTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAG 4021
             |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||:||
Db        1196 GTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAG 1255

Qy       4022 CTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAAT 4081
             ||:|||||||||||||:|||:| ||| |||||||| ||||||||:|||||||:|||| 
Db        1256 CTGGCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAAT 1315

Qy       4082 ATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARA 4141
             |||||||||||||||:|:| ||||:|||||| ||||||||||::|||||||:|:|
Db        1316 ATATATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGA 1375

Qy       4142 GAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARA 4201
             |||||||||||||||||||:|||:||||||||||||:|||||||||||||||||||||:|
Db        1376 GAAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAA 1435

Qy       4202 TTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCAN 4261
             ||||||:||||:|||||||||||||:|||||||:||| |||||||||:|||||||||||
Db        1436 TTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAG 1495

Qy       4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
```

```
                ||||||||:||||||||||:|||:|| :  |||||  :|||||||| :||||||||
Db       1496 AAAGGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCC 1555

Qy       4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
                ||||||  |||||||||||||||||||||||||    |   ||| ||||||  ||||||||||
Db       1556 ACAAGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCC 1615

Qy       4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
                ||||||||  |||||||||||||||||||||||||||||||| |||||
Db       1616 CTCCGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAG 1660


RESULT 64
AAS84201
ID   AAS84201 standard; cDNA; 1660 BP.
XX
AC   AAS84201;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20005.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG20014.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 20005; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1660 BP; 511 A; 396 C; 351 G; 402 T; 0 U; 0 Other;

  Query Match             17.2%;  Score 1285.7;  DB 1;  Length 1660;
  Best Local Similarity   77.2%;  Pred. No. 0;
  Matches 1470;  Conservative   92;  Mismatches   98;  Indels  245;  Gaps    6;

Qy       2522 GCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAA 2581
                ||||||||||| |||||||||||||||||||| ||||||| | |||||||||| ||||
Db          1 GCAGTACCCCCTTAGACCCAAGGCCCAACAAGGACTCCAAAAAACTGTTAAGGACCTAAA 60

Qy       2582 AGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGG 2641
                ||||||||||  |||||||||||||||   |  |||||||  || | |         | ||  |
Db         61 AGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAATACTCC-------AACTGTAG 113

Qy       2642 AGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGA 2701
                  | ||||||||||  ||||||||| |||||||||| |||||||||||||||||| ||
Db         114 GAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAAATCTCAGGATTATCAAT-GA 171

Qy       2702 GGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACC 2761
                | | ||||| | | |||||||||||| ||||| |||||||||| ||::||||||||||||
Db         172 GACTGTTGTTCCTCTATACCCAGCTGAACCTAACCCTTATACTCTGCTTTCCCAAATACC 231

Qy       2762 AGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGT 2821
                |||||||||||||||||  || |:|||||||||||:|||||||||||||||||||||||||
Db         232 AGAGGAAGCAGAGTGGTATATAGTCCTGGACCTTAAGGATGCCTTCTTCTGCATCCCTGT 291

Qy       2822 ACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACT 2881
                |||||||| || |||||| |||||||||||||||| |||||||| ||:| ||||||||
Db         292 ACATCCTGATTCCCAATTCCTGTTTGCCTTTGAAGATCCTTCAAACTCAACGTCTCAACT 351

Qy       2882 CACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATT 2941
                ||||||||||:|||||| ||||||||||||||||||:|||||||||||||||||||||||
Db         352 CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATT 411

Qy       2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
                |||||||||  |||||:|||:|::|||||||||||||||||||||||| :|||:|||||||||
Db         412 AGCCCAAGATTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGA 471

Qy       3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
                |||||||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:||
Db         472 TTTACTTTTAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAA 531

Qy       3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
                ||||||| |:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||
Db         532 TTTCCTCACCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCA 591

Qy       3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
                |||||||||||||||:||||:|||||||||||||    |:|||||||||||||||:::||
Db         592 GGTTAAATACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACAT 651

Qy       3182 CCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGG 3241
                |||||||||||||||||||||||||||||||||||||||||||||||:|:::|||||||||
Db         652 CCAGCCTATACTGGCTTATCCTCTTCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGG 711

Qy       3242 CRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAW 3301
                |:|||:|||:|||| |||||: |||||| |||||||:  ||:|||||| ||| :|||||:
Db         712 CATAACAGGTTTCTACCGAATGTGGATT-CCCAGGTACAGCAAAATAGGCAGATCATTAT 770

Qy       3302 ATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAG 3361
                |||||:|||||| ||||||||||||||||||||||||||
Db         771 ATACACTAATTAAAGAAACTCAGAAAGCCAATACCCATTTA---------------------- 811

Qy       3362 YMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTT 3421

Db         812 ---------------------------------------------------------- 811

Qy       3422 GCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAG 3481

Db         812 ---------------------------------------------------------- 811

Qy       3482 TCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTG 3541

Db         812 ---------------------------------------------------------- 811
```

```
Qy       3542 ATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAG 3601
                                                 ||||||||||||||||||||::|||
Db        812 ----------------------------------GTAGTGGTGGCAGTAGCAGTCTTAG 836

Qy       3602 TATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKG 3661
              | |||||||  ||||||| ||||||||:||||||||||||||||||||||||||||:|
Db        837 TGTCTGAAGTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATG 896

Qy       3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
              ||||:: |||||||||||||||||||||||||||||||||||||:||||||||||||
Db        897 TGAATGCCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAAT 956

Qy       3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
              :|||||||||||||||||:|||||||||||:|||| |||||||||| ||||||||||
Db        957 ATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACC 1016

Qy       3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTT 3841
              |||:| |||||||||||||||||||||||||:||:||||||| ||||||:|||||
Db       1017 CAGTCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTG 1076

Qy       3842 CTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTC 3901
              |||||||||||||| |||||||||| ||:||||:||||||:||||||||||: ||| ||
Db       1077 CTCAAACCTATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TC 1135

Qy       3902 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
              |||||||||||||||||||||||||||||||||||||||||||||| |||||:||||||||
Db       1136 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCA 1195

Qy       3962 GTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAG 4021
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||:||
Db       1196 GTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAG 1255

Qy       4022 CTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAAT 4081
              ||:|||||||||||||:|||:| ||| |||||||| ||||||||:||||||||:||||
Db       1256 CTGGCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAAT 1315

Qy       4082 ATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARA 4141
              ||||||||||||||:|:| ||||:|||||||| |||||||||||::|||||||:|:|
Db       1316 ATATATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGA 1375

Qy       4142 GAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARA 4201
              |||||||||||||||||||:|||:||||||||||||:|||||||||||||||||||:|
Db       1376 GAAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAA 1435

Qy       4202 TTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCAN 4261
              ||||||:||||:|||||||||||||:||||||||:||| |||||||||:|||||||||
Db       1436 TTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAG 1495

Qy       4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
              |||||||||:|||||||||||:|||:|| :  |||||  :|||||||| :||||||||
Db       1496 AAAGGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCC 1555

Qy       4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
              ||||||| |||||||||||||||||||||||||    |   ||| |||||| ||||||||||
Db       1556 ACAAGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCC 1615

Qy       4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
               |||||||| |||||||||||||||||||||||||||| ||||| |
Db       1616 CTCCGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAG 1660
```

RESULT 65
AAX25655
ID AAX25655 standard; cDNA to mRNA; 1321 BP.
XX
AC AAX25655;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.6A2.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97FR-00008815.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
DR WPI; 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
PS Claim 1; Page 48-49; 106pp; French.
XX
CC This sequence represents clone cl.6A2 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility
XX
SQ Sequence 1321 BP; 365 A; 352 C; 277 G; 327 T; 0 U; 0 Other;

Query Match 17.0%; Score 1270.4; DB 1; Length 1321;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1291; Conservative 22; Mismatches 7; Indels 4; Gaps 3;

```
Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
```

```
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              ||||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              |||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:|||||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              |||| |
Db       1317 GAGG 1320


RESULT 66
AAA59205
ID   AAA59205 standard; DNA; 1321 BP.
XX
AC   AAA59205;
XX
DT   07-NOV-2000  (first entry)
XX
DE   5' non coding and partial gag fragment of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 42-43; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 1321 BP; 365 A; 352 C; 277 G; 327 T; 0 U; 0 Other;

  Query Match             17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
```

```
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              ||||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:|||||||||||||||||||||||||||||||||||     ||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              |||
Db       1317 GAGG 1320


RESULT 67
AAI19284
ID   AAI19284 standard; DNA; 1472 BP.
XX
AC   AAI19284;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #9217 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 9217; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1472 BP; 481 A; 335 C; 327 G; 329 T; 0 U; 0 Other;

  Query Match            16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity  87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||| | |||| ||||||||||||:||||:||||||||||||||||:
Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              |||||||||||||||:::|||||||| ||||||||||||||||||:||||||||||||
Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::||| |||||:|| :|||:||||||||||:||||||| |||||||: |
Db         121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:|||||| |||:|||||:|||||:|||| ||||||||||||||||||||||||||||
Db         180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:|||||||||:|:||||||:::||| |||||||||||||  ||  | | |||  |||
Db         240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              |||||:|||||||    ||||:||||||||:|||||||||||| :||:|||::||||||
Db         298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||||||||:||:||||||:||||||||||||||:|||:|||||: ||| |||||||::
Db         357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::||||||:||||||||:|||||||||||||||||||:|:||||||||||: ||
```

```
Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637

Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697

Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757

Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817

Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877

Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937

Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997

Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057

Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117

Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177

Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237

Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297

Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357

Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417

Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474

Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534

Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574

Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472
```

RESULT 68
ABA64295

```
ID   ABA64295 standard; DNA; 1472 BP.
XX
AC   ABA64295;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #12600.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 12600; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1472 BP; 481 A; 335 C; 327 G; 329 T; 0 U; 0 Other;
```

Query Match 16.4%; Score 1229.2; DB 1; Length 1472;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 1293; Conservative 96; Mismatches 77; Indels 14; Gaps 7;

```
Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160

Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220

Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280

Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340

Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400

Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
```

```
Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517

Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577

Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637

Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697

Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757

Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817

Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877

Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937

Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997

Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057

Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117

Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177

Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237

Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297

Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357

Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417

Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474

Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534

Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574

Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472


RESULT 69
AAI44454
ID   AAI44454 standard; DNA; 1472 BP.
XX
AC   AAI44454;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #13140 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 13140; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 1472 BP; 481 A; 335 C; 327 G; 329 T; 0 U; 0 Other;

  Query Match            16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160

Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220

Db          61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280

Db         121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
```

```
Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400

Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457

Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517

Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577

Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637

Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697

Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757

Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817

Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877

Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937

Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997

Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057

Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117

Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177

Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237

Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297

Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357

Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417

Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474

Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534

Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574

Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472
```

RESULT 70
ABS38070
ID ABS38070 standard; DNA; 1472 BP.
XX
AC ABS38070;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 13060.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 1; SEQ ID NO 13060; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1472 BP; 481 A; 335 C; 327 G; 329 T; 0 U; 0 Other;

Query Match 16.4%; Score 1229.2; DB 1; Length 1472;

```
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||||  | ||||  |||||||||||||:||||:|||||||||||||||:
Db             1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              ||||||||||||||||:::|||||||| ||||||||||||||||||:||||||||||||
Db            61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::||| |||||:|| :|||:||||||||||:||||||| |||||||: |
Db           121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:|||||| |||:|||||:|||||:||||| ||||||||||||||||||||||||||||
Db           180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:|||||||||:|:||||||:::||| |||||||||||||   ||  | | |||  |||
Db           240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              |||||:||||||||   ||||:||||||||:|||||||||||| :||:|||::||||||
Db           298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||||||||:||:||||||:|||||||||||||||||:|||:|||||: ||| |||||||::
Db           357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::|||||||:||||||||:|||||||||||||||||||:|:|||||||||| ||
Db           417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
              |||||| ||||||||||||::|||||||||||||||||||| |||||||| :|||||||
Db           477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
              ||||||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db           537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
              ||||||||||:||||||||||:|||||||||| | ||||||:| ||||||||||:||
Db           597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
              |||||||||||||||||||||||||:| |||||||||||| ||||||||||||||||
Db           657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
              ||:|||||||||||||:|||||  |||||||||| ||||||| ||||||||||:||||:|
Db           717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
              ||||: |||||||||:|||| ||||||||||||||||||||||||| |||||||||
Db           777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
              || |||| |||||:|||||||||||| ||||||||||||||||:||||||||||||||
Db           836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
              ||||||||||||||||||||:|||||:|||||||||||||||:|||:|  |||| |||||
Db           896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
              ||   |||||:||||||||:|||||||||||||||||||||:|:||||||:||||||| ||
Db           956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
              ||||||||||::||||||||:|:|||||||| ||||||||||||:|||:||||||||||||
Db           1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
              |:|||||||||||||||||||:|||||||:|||| ||||||||||||:|||||||||:|
Db           1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
              ||||||||||||:|||||||||| |||||||||:||||||||||:|||:|||:||||||
Db           1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
              || :| ||||||||:|||||||||||||||||||||||||||||||||||||||||||  ||
Db           1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
                  ||| ||||||  |||||||||| || ||||||||||||||||||||||  || ||||
Db           1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
              | |||||||||||||||| ||||    ||||| |||||||| ||| || |||||||  ||
Db           1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
              ||||| |||||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db           1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
              |||||||||||||||  |||| ||||||| |:||| |||
Db           1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472
```

```
RESULT 71
ABS12550
ID   ABS12550 standard; DNA; 1472 BP.
XX
AC   ABS12550;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 12541.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
```

PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 1; SEQ ID NO 12541; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1472 BP; 481 A; 335 C; 327 G; 329 T; 0 U; 0 Other;

Query Match 16.4%; Score 1229.2; DB 1; Length 1472;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 1293; Conservative 96; Mismatches 77; Indels 14; Gaps 7;

```
Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
Db        180 GCGAAATAGCTAGGCCATTATATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432
```

```
Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
              ||||||||||||||||  |||| |||||||  |:||| |||
Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472


RESULT 72
ABN97928
ID   ABN97928 standard; DNA; 1326 BP.
XX
AC   ABN97928;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Human retroviral sequence gag.
XX
KW   Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW   multiple sclerosis; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9967395-A1.
XX
PD   29-DEC-1999.
XX
PF   23-JUN-1999;   99WO-FR001513.
XX
PR   23-JUN-1998;   98FR-00007920.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Alliel PM,  Perin J,  Rieger F;
XX
DR   WPI; 2000-160587/14.
XX
PT   New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT   for diagnosis, treatment and prevention of autoimmune and neurological
PT   diseases.
XX
PS   Claim 2; Page 129; 225pp; French.
XX
CC   The present invention relates to new nucleic acid sequences of human
CC   endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC   Regulatory elements associated with HERV-7q may alter expression of other
CC   genes (even remote genes) on the same chromosome, inducing immunological
CC   and/or neurological changes (which may be pathological or protective/
CC   curative). HERV-7q peptides can be used to improve efficiency of the
CC   immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC   sequences can be used in immunogenic or vaccinating compositions, for
CC   protection against autoimmune diseases, particularly multiple sclerosis.
CC   The peptides may also be used (by sequence comparison) to detect/identify
CC   endogenous retroviruses that are abnormally expressed in cancer,
CC   neuropathologies or other autoimmune diseases. The present sequence was
CC   used to illustrate the invention
XX
SQ   Sequence 1326 BP; 425 A; 306 C; 303 G; 292 T; 0 U; 0 Other;

  Query Match             16.0%;  Score 1201.4;  DB 1;  Length 1326;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1295;  Conservative   16;  Mismatches   14;  Indels   12;  Gaps    9;

Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
              ||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
              |:||||::|||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
              ||||||||||:|||||||||||||||||||||||||||||||||||     ||||||||
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
              |||| |||||||| ||||||||||||||||||||:|||||||||||||||:|||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
              |||||   ||| | |     ||||| ||||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711

Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
              |||||:|||||||||||||||||:|||| ||:||||:|||||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009

Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
              |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
```

```
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
              |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326


RESULT 73
AAV43228
ID   AAV43228 standard; cDNA; 1511 BP.
XX
AC   AAV43228;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus fragment 8.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2. .1060
FT                   /*tag=  a
FT                   /product= "Encodes protein AAW71070"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71071.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 191-192; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) genomic fragment used in the method of the invention.
CC   The invention provides complete or partial genomic sequences of the MSRV-
CC   1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC   genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 1511 BP; 481 A; 361 C; 341 G; 328 T; 0 U; 0 Other;

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||||:|||||||:|||| |||| |||||||  ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
```

```
                ||||||||||||||||||||||||||| |||||||||||||||:|||| |||||||
Db       121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy       786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
               | |||||||||||||||||||||||| || |||||||||||||||||||||||||||||
Db       181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy       844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
             |||||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db       241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy       904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
               |||||     | |||||||||||||||| ||::|||||||||||||||||| ||||||||
Db       301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy       959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
             |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db       361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||| ||| |
Db       421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db       480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||||| ||||||||||||||||||||||
Db       540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db       600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |  |||||||||
Db       659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||       |||||||||||||:|||||||||| ||:|||||||||
Db       717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db       773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:|||||| |||||||||:|||| ||: |||:|||||||||||||||||||||| |
Db       829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db       888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||| ||||||||||||||||||||| ||||
Db       947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||| |||||||        | |||| |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||  |||||||||||||||| ||||||||||||||||||||||||
```

```
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
             |||||||| |||||||||| | |||||||||||||   |||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
             |||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
             |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
             |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 74
AAX29706
ID   AAX29706 standard; DNA; 1511 BP.
XX
AC   AAX29706;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Clone CL2 from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis; ss.
XX
OS   Viruses.
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   P-PSDB; AAW99556.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 1; Page 42-43; 83pp; French.
XX
CC   This sequence represents clone CL2 from a novel multiple sclerosis
CC   related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC   prophylactic or therapeutic compositions to inhibit expression of a
CC   multiple sclerosis related virus and/or virus associated with rheumatoid
CC   polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC   17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 1511 BP; 481 A; 361 C; 341 G; 328 T; 0 U; 0 Other;

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||

Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || |||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||||||||| ||||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||||| |||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| |||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| |||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||||||| |||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| ||||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |   |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      |||||||||||||:|||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| ||||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||| ||||||||||||||||||||   ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126
```

```
Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
             |||||||||||| ||| |||||||||||||| ||||||||       | ||||  ||||||| 
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
             ||||  ||||||||||||  ||||||||||||||||| ||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| 
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
             |||||||| |||||||||||  | ||||||||||||||  |||||||||||||||||||| 
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
             ||||||||||||||||  ||| |  | |||||||||||||| |||||||||||||||||| 
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
             |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||||| | 
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
             |||||| |||||||||||| |||| 
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

RESULT 75
AAA59188
ID  AAA59188 standard; DNA; 1511 BP.
XX
AC  AAA59188;
XX
DT  07-NOV-2000  (first entry)
XX
DE  Endogenetic retroviral fragment integrated into the human genome.
XX
KW  Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW  gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS  Homo sapiens.
XX
PN  WO200043521-A2.
XX
PD  27-JUL-2000.
XX
PF  21-JAN-2000; 2000WO-FR000144.
XX
PR  21-JAN-1999;  99FR-00000888.
XX
PA  (INMR ) BIO MERIEUX.
XX
PI  Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR  WPI; 2000-499229/44.
XX
PT  New nucleic acid from human endogenous retrovirus, useful e.g. for
PT  diagnosis of autoimmune disease and complications of pregnancy, contains
PT  at least part of the gag gene.
XX
PS  Claim 4; Page 35; 53pp; French.
XX
CC  The present sequence represents an endogenetic retroviral nucleic acid
CC  fragment, which is associated with an autoimmune disease, and is
CC  integrated into the human genome. The fragment is originally derived from
CC  a novel retrovirus, human endogenous retrovirus W (HERV-W). The fragment
CC  comprises (or consists of) at least part of the gag gene of the HERV-W
CC  retrovirus, which is associated with either autoimmune disease, failure
CC  of pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC  proteins derived from it, are useful for diagnosis of autoimmune disease
CC  (specifically multiple sclerosis) and for monitoring pregnancy. The
CC  nucleic acid fragments may also be used for in situ labelling of isolated
CC  chromosomes, while the transcription product can be used to study or
CC  monitor T cell proliferation in vitro
XX
SQ  Sequence 1511 BP; 481 A; 361 C; 341 G; 328 T; 0 U; 0 Other;

```
  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| |||||||| ||||||||||||| 
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||||| |||||||||                             
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||||:|||| |||||||  
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||| || |||||||||||||||||||||||||||| 
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||||| |||||||| ||||||||||| 
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||| ||::|||||||||||||||||| |||||||| 
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||| 
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||||| ||| | 
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:| 
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||| |||||||||||||||||||||| 
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||| |  |||||  
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||| |||   |   |||||||| 
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||       |||||||||||||:|||||||||||| ||:|||||||| 
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |      ||||| ||||||||||||  || |||||||||||||| || 
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| ||||||||||:|||| ||: |||:|||||||||||||||||||| | 
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||||| |||||| |||||||| |||||||||||||||| 
```

```
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||||  ||||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||||  ||| ||||||||||||| ||||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| ||||||||||||  |||||||||||||||| ||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||  ||||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | ||||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

```
RESULT 76
ADG14858
ID   ADG14858 standard; cDNA; 1511 BP.
XX
AC   ADG14858;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #19.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 117; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1511 BP; 481 A; 361 C; 341 G; 328 T; 0 U; 0 Other;
```

Query Match 14.9%; Score 1117.7; DB 1; Length 1511;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 1327; Conservative 18; Mismatches 118; Indels 61; Gaps 12;

```
Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| ||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               ||||||||||||||||||||||||||||| |||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||||| ||||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||| |||   |  |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||       |||||||||||||:|||||||||||| ||:|||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
```

```
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| ||||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| |||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||||  ||||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||||| ||| |||||||||||||| |||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| ||||||||||||  |||||||||||||||| |||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||  |||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | |||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

```
RESULT 77
AAX25660
ID   AAX25660 standard; cDNA to mRNA; 1136 BP.
XX
AC   AAX25660;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W clone cl.C4C5.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Claim 1; Page 59-60; 106pp; French.
XX
CC   This sequence represents clone cl.C4C5 of the human endogenous retrovirus
CC   (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC   by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC   rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC   dependent diabetes and related pathologies) and of abnormal or
CC   unsuccessful pregnancy and can be used as chromosomal markers for
CC   susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||||:|||||||||||||||||:|||||||||||:||||||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:||||||||||||||||||:||||||:||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660
```

```
Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:|||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116
```

```
RESULT 78
AAA59210
ID   AAA59210 standard; DNA; 1136 BP.
XX
AC   AAA59210;
XX
DT   07-NOV-2000  (first entry)
XX
DE   3' pol gene and 3' non coding sequences of HERV-W from human genome.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 46; 53pp; French.
XX
CC   The present sequence represents an endogenetic retroviral nucleic acid
CC   fragment, which is associated with an autoimmune disease, and is
CC   integrated into the human genome. The fragment is originally derived from
CC   a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC   retrovirus is associated with autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;
```

```
  Query Match           14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity  97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             |||||||||||||||:||||||||||||||||||:||||||||||:||||||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:||||||||:||||||||||||||||||:||||||:||||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:|||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900
```

```
Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:|||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 79
ADB84412
ID   ADB84412 standard; DNA; 1511 BP.
XX
AC   ADB84412;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #29.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Disclosure; Page 83-84; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 1511 BP; 471 A; 350 C; 332 G; 325 T; 0 U; 33 Other;

  Query Match             14.6%;  Score 1093.5;  DB 1;  Length 1511;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches 1301;  Conservative   17;  Mismatches  145;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:|||||||:|||| |||| ||||||| |||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTN 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAN 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || |||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTN 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| |||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTN 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||      | |||||||||||||||| ||::|||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAN 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCN 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              ||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db        480 NTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              ||| ||||||||||:|||||||||||||||||||||| ||||||||||||||||||||
Db        540 NTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              |||| |||| ||| |||||||||||||||||||||||||||||||| |  |||||
Db        600 NTATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              | |||| |||:|||||||||||| |||||||||||||||| |||   |  |||||||||
Db        659 GNAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||  | ||  | ||       |||||||||||||:|||||||||||| ||:||||||||
Db        717 CAGNGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
```

```
Db        773 TGGG---ACANAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
Db        829 GCTAGAAGGACNGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
Db        888 TAACACA-GGGAANGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
Db        947 AGGAAGCATACCTNCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
Db       1007 TTTATCACTCAGNNAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
Db       1067 AGCAGAACTTAGNNACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
Db       1127 AGGAGCAGGCGANNCGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
Db       1187 CATGGCCCTCAGNNAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
Db       1247 AATGCCTAATAGNNCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
Db       1307 AAGTAGAAATAANNCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
Db       1367 CCACTGCCCCAGNNGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
Db       1427 GCAGGACTGAGGNNGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
Db       1487 GGTATGTTTGACNNTTGAGAGCCA 1510
```

```
RESULT 80
ABN97949
ID   ABN97949 standard; DNA; 1080 BP.
XX
AC   ABN97949;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Human retroviral HERV-7q gag coding sequence.
XX
KW   Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW   multiple sclerosis; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9967395-A1.
XX
PD   29-DEC-1999.
XX
PF   23-JUN-1999;   99WO-FR001513.
XX
PR   23-JUN-1998;   98FR-00007920.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Alliel PM,  Perin J,  Rieger F;
XX
DR   WPI; 2000-160587/14.
XX
PT   New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT   for diagnosis, treatment and prevention of autoimmune and neurological
PT   diseases.
XX
PS   Claim 3; Page 148-150; 225pp; French.
XX
CC   The present invention relates to new nucleic acid sequences of human
CC   endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC   Regulatory elements associated with HERV-7q may alter expression of other
CC   genes (even remote genes) on the same chromosome, inducing immunological
CC   and/or neurological changes (which may be pathological or protective/
CC   curative). HERV-7q peptides can be used to improve efficiency of the
CC   immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC   sequences can be used in immunogenic or vaccinating compositions, for
CC   protection against autoimmune diseases, particularly multiple sclerosis.
CC   The peptides may also be used (by sequence comparison) to detect/identify
CC   endogenous retroviruses that are abnormally expressed in cancer,
CC   neuropathologies or other autoimmune diseases. The present sequence was
CC   used to illustrate the invention
XX
SQ   Sequence 1080 BP; 358 A; 239 C; 240 G; 243 T; 0 U; 0 Other;
```

```
  Query Match             12.8%;  Score 960.2;  DB 1;  Length 1080;
  Best Local Similarity   96.4%;  Pred. No. 0;
  Matches 1052;  Conservative   16;  Mismatches   11;  Indels   12;  Gaps    9;

Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595
```

```
Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
             ||||||||||||    ||| | |      ||||| ||||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
             ||||||||||||:||||||||||||||||:|||| ||:||||:|||||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
             ||||||||||||
Db       1070 TGAATGCCTAA 1080
```

RESULT 81
AAT96476
ID AAT96476 standard; cDNA; 1167 BP.
XX
AC AAT96476;
XX
DT 26-FEB-1998 (first entry)
XX
DE MSRV-1 clone FP6.
XX
KW Multiple sclerosis related virus 1; env; pol; gag; RT-PCR; primer; ss;
KW amplification; reverse transcriptase; rheumatoid arthritis; diagnosis.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN WO9706260-A1.
XX
PD 20-FEB-1997.
XX
PF 02-AUG-1996; 96WO-FR001244.
XX
PR 03-AUG-1995; 95FR-00009643.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX
DR WPI; 1997-154266/14.
XX
PT New viral material and nucleotide fragments associated with multiple
PT sclerosis and rheumatoid arthritis - also related peptide(s) and
PT antibodies, used for diagnosis, treatment and as vaccines.
XX
PS Claim 1; Page 112-113; 188pp; French.
XX
CC Sequences AAT96468-77 represent claimed nucleic acid sequences derived
CC from the multiple sclerosis related virus 1 (MSRV-1), especially from the
CC env, pol and part of the gag gene region, but exclude the sub-region
CC represented by AAT96527. This sequence is from clone FP6 and is
CC especially from the pol gene region. The sequence was isolated PCR using
CC primers AAT96558-60. The above sequences are associated with multiple
CC sclerosis and some are also associated with rheumatoid arthritis. Nucleic
CC acid sequences from this virus can be used in the diagnosis of multiple
CC sclerosis and rheumatoid arthritis
XX
SQ Sequence 1167 BP; 384 A; 254 C; 270 G; 258 T; 0 U; 1 Other;

Query Match 12.7%; Score 952.8; DB 1; Length 1167;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1013; Conservative 77; Mismatches 73; Indels 10; Gaps 6;

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
             ||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::|||  |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
             |||||||||  ||  ||  |||  ||||||||:|||||||||||:||||||||:||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::|||||  |||||||||:||:|||||||:||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: |||||||||||::||||::|||||:||||||||:||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:|||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             ||||||||||||||||:|||||||||||||||||||||||||||||:|||||::||||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||||  |||||||||||||||||||: ||||||||||||: |||| ||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: ||||||||||:|||| || ||||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||||:||:|||||||||||||||:|||||  ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  |||||||||||||:||||:|||||:||||||||||||:|||| ||||||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||||| ||||||:|||||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||||||||||||||||||||||||:| ||:|||||||||||
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:|:|||||||||||||||||||||:||||||||:|||||||||| ||||
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| |||||||||||::|||||||:|:||| ||||||||
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834
```

```
Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||:|||:||||||||||||: ||||| |||||||||||||:|||||||:|||
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:||||||||:|||||||||||||: |||||||||  ||| | ||:
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || ||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||| ||| ||||| || |||||||| |||     | ||| |
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             |||||||  || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167
```

RESULT 82
AAV43170
ID AAV43170 standard; cDNA; 1167 BP.
XX
AC AAV43170;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus clone FP6 DNA sequence.
XX
KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN WO9823755-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-IB001482.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX
DR WPI; 1998-322732/28.
XX
PT New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT treatment of, e.g. multiple sclerosis.
XX
PS Example 14; Fig 38; 286pp; English.
XX
CC The present sequence represents the multiple sclerosis (MS) associated
CC retrovirus (MSRV) clone FP6 DNA sequence containing a portion of the pol
CC gene. The invention provides complete or partial genomic sequences of the
CC MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC genes. The invention also provides antibodies raised against the
CC polypeptides. The genomic sequences, polypeptides and antibodies are also
CC claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC associated viruses, and also for the prevention and treatment of
CC infection with these viruses
XX
SQ Sequence 1167 BP; 384 A; 254 C; 270 G; 258 T; 0 U; 1 Other;

Query Match 12.7%; Score 952.8; DB 1; Length 1167;

Best Local Similarity 86.4%; Pred. No. 0;
Matches 1013; Conservative 77; Mismatches 73; Indels 10; Gaps 6;

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
             ||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::||| |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
             |||||||||  ||  ||  |||  ||||||||:||||||||||||:||||||||:|||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::|||||| ||||||||:||:||||||:|||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: ||||||||||::||||::|||||:|||||||:|||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             |||||||||||||||||:|||||||||||||||||||||||||||||:|||||::|||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||| ||||||||||||||||||||||: ||||||||||: |||| ||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: |||||||||||:||||  || ||||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||:||:|||||||||||||:||||| ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  |||||||||||||||:|||||:|||||:||||||||||||:|||| ||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||| |||||:|||||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||| ||||||||||||||||||||:| ||:||||||||
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:|:|||||||||||||||||||:||||||||:|||||||||| ||||
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| ||||||||||::|||||||:|:||| ||||||||
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||:|||:|||||||||||: ||||| |||||||||||||:|||||||:|||
Db         835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:||||||||:|||||||||||||: |||||||||  ||| | ||:
Db         895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db         955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || ||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
```

```
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             ||||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             ||||||| || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167
```

RESULT 83
ADB84352
ID ADB84352 standard; cDNA; 1167 BP.
XX
AC ADB84352;
XX
DT 04-DEC-2003 (first entry)
XX
DE MSRV-1 pol/RT region clone FB6.
XX
KW MSRV; ss; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003039664-A1.
XX
PD 27-FEB-2003.
XX
PF 26-NOV-1997; 97US-00979847.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
DR P-PSDB; ADB84524, ADB84525, ADB84526.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Example 14; Fig 38; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is an MSRV cDNA clone. Note: The SEQ ID numbers for the
CC sequences as displayed in the main body of the patent do not match the
CC SEQ ID numbers in the sequence listing. Consequently those sequences
CC mentioned in the claims may not be the sequences the authors intended to
CC claim.
XX
SQ Sequence 1167 BP; 384 A; 254 C; 270 G; 258 T; 0 U; 1 Other;

Query Match 12.7%; Score 952.8; DB 1; Length 1167;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1013; Conservative 77; Mismatches 73; Indels 10; Gaps 6;

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
             ||||||||||||||||||||||||||||||||||:||||||||:|:| ||||:::||| |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
             ||||||||||  ||  ||  |||  ||||||||:||||||||||||:||||||||:||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::||||||  ||||||||:||:||||||:||||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: |||||||||||::|||||::||||||:|||||||:||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             |||||||||||||||||:|||||||||||||||||||||||||||||:|||||::|||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||||  |||||||||||||||||||||: |||||||||||: |||| ||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: |||||||||||:|||| || |||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||:||:|||||||||||||||:|||||  ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTATAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||| |||||||:|||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||| ||||||||||||||||||||:| ||:||||||||
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:|:|||||||||||||||||||:||||||||:|||||||||| ||||
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||:|||||||| ||||||||||::|||||||:|:||| |||||||||
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||:|||:|||||||||||||: ||||| |||||||||||||:||||||:|||
```

```
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:||||||||||||:||||||||:||||||||||||: |||||||||  ||| | ||:
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || |||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             |||||||  || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 84
ADG14798
ID   ADG14798 standard; cDNA; 1167 BP.
XX
AC   ADG14798;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1 pol ORF #1.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 10; SEQ ID NO 57; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1167 BP; 384 A; 254 C; 270 G; 258 T; 0 U; 1 Other;

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
             |||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::||| |||
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
             |||||||||  ||  ||  |||  ||||||||:||||||||||||:||||||||:||||
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::|||||||  ||||||||:||:|||||||:|||||||||||||||
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: ||||||||||||::||||::|||||||:||||||||:|||||||| |||||
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:|||||||||||: ||||| ||  ||||||||||||::|||| |||||| ||
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             |||||||||||||||||:||||||||||||||||||||||||||||||:|||||::|||||
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||||  |||||||||||||||||||||: |||||||||||: |||| ||||
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: ||||||||||:||||   || |||||||||||||||||||:| |||||
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||:||:|||||||||||||||:||||| ||||||||||
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  |||||||||||||:||||:|||||:|||||||||||:|||| |||||||||||||
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||| |||||||:|||||||||||||| |||||||
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             |||||||||:|||||||||||| ||||||||||||||||||||:| ||:|||||||||
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:|:|||||||||||||||||||||||:||||||||:|||||||| ||||
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:||||||:|||||||| ||||||||||::||||||||:|:||| ||||||||
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||:|||:|||||||||||||: ||||| |||||||||||||||:|||||||:|||
Db         835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:||||||||||||:||||||||:||||||||||||: |||||||||  ||| | ||:
Db         895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db         955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || |||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db         1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
```

```
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
              ||||||| || || |||||||||| ||||| |
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 85
AAI28063
ID   AAI28063 standard; DNA; 1117 BP.
XX
AC   AAI28063;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #17996 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 17996; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match           12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
              |||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
              ||||||||  ||  | | |||  ||||||||:||||||||   ||||:||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
              ||||||||| :||:|||::|||||| |||||||||:||:||||||:|||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
              |:|||:|||||: ||| |||||||::||||::||||||:|||||||:|||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
              |||||||||:|:||||||||||: |||||||| |||||||||||||::|||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
              ||||||| ||||||||| :|||||||||||||||||||||||||||||:|||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
              |||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
              | | ||||||:| ||||||||||:|||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
              |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||| ||||||||||:||||:|||||: ||||||||||:|||| ||||||||||||
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
              |||||||||||||||| |||||||||| |||| |||||:|||||||||||| |||||||
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
              ||||||||||:|||||||||||||||||||||||||||||||||||:||||:|||||||
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
              ||||||||:|||:|  |||| |||||||   |||||:||||||:|||||||||||||||
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
              |||||:|:||||||:|||||||| ||||||||||::||||||||:|:|||||||| ||
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
              ||||||||||:|||:|||||||||||:|||||||||||||||||||:|||||||:|||
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
              | ||||||||||||:|||||||:||||||||||||:|||||||||| |||||||||:
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
              ||||||||||:|||:|||:|||||||| :| ||||||||:|||||||||||||||||||
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
              |||||||||||||||||||||  ||   ||| ||||||  |||||||| || ||||||
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
              |||||||||||||||  || ||||| ||||||||||||||||
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 86
ABA76381
ID   ABA76381 standard; DNA; 1117 BP.
XX
AC   ABA76381;
```

```
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #24686.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 24686; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match           12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:|||||||||:|:|||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  ||||||||:||||||||   ||||:|||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             |||||||||  :||:|||::|||||| ||||||||||:||:||||||:|||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::||||::||||||:||||||||:||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:|||||||||: |||||||| ||||||||||||::|||||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| ||||||||| :|||||||||||||||||||||||||||||:||||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||||
Db         358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | ||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db         418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db         478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             |||||||  ||||||||||:||||:|||||: |||||||||||:|||| ||||||||||||
Db         538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:|||||||||||| |||||||
Db         597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             ||||||||||:|||||||||||||||||||||||||||||||||||:||||:|||||||
Db         657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   |||||:|||||||:|||||||||||||||
Db         717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:||||||:|||||||| |||||||||||::|||||||:|:|||||||| ||
Db         774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:||||||||||||:|||||||||||||||||||||:|||||||:|||
Db         834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | |||||||||||||:|||||||:|||||||||||||:|||||||||| |||||||||:
Db         894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|||:|||||||| :| ||||||||:|||||||||||||||||||||
Db         954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||||  ||   ||| ||||||  |||||||| || ||||||
Db        1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
             |||||||||||||||||  || ||||| ||||||||||||||||
Db        1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117
```

```
RESULT 87
AAI57054
ID   AAI57054 standard; DNA; 1117 BP.
XX
AC   AAI57054;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #25740 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
```

```
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 25740; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 88
ABS50596
ID   ABS50596 standard; DNA; 1117 BP.
XX
AC   ABS50596;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 25586.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
```

```
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 25586; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  ||||||||:||||||||   ||||:|||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             ||||||||| :||:|||:|||||| |||||||||:||:||||||:|||||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::||||::|||||||:||||||||:||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:|||||||||: |||||||| ||||||||||||::|||||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| ||||||||| :|||||||||||||||||||||||||||||||:|||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | |||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||| |||||||||:||||:|||||: ||||||||||:|||| ||||||||||||
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:|||||||||||||| |||||||
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             ||||||||||:|||||||||||||||||||||||||||||||||:||||:|||||||||
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   ||||||:||||||||:|||||||||||||
Db         717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:||||||:|||||||| ||||||||||||::|||||||:|:|||||||| ||
Db         774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:|||||||||||:|||||||||||||||||||:|||||||:|||
Db         834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | |||||||||||||:|||||||:|||||||||||||:|||||||||| |||||||||:
Db         894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|||:|||||||| :| ||||||||:||||||||||||||||||||
Db         954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||||  ||   ||| ||||||  |||||||| || ||||||
Db        1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
             |||||||||||||||||  || ||||| ||||||||||||||||
Db        1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 89
ABS24539
ID   ABS24539 standard; DNA; 1117 BP.
XX
AC   ABS24539;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 24530.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
```

```
XX
PS   Claim 4; SEQ ID NO 24530; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:||||||||:|:|||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  ||||||||:||||||||   ||||:|||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             |||||||||  :||:|||::|||||| |||||||||:||:||||||:|||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::||||::||||||:|||||||:||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:||||||||||: |||||||| ||||||||||||::||||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| ||||||||| :||||||||||||||||||||||||||||:||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | |||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             ||||||| ||||||||||:||||:|||||: ||||||||||:|||| ||||||||||||
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:||||||||||| |||||||
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             ||||||||||:|||||||||||||||||||||||||||||||||:||||:|||||||
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   |||||:||||||||:|||||||||||||
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:||||||:|||||||| ||||||||||::||||||||:|:|||||||  ||
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:||||||||||||:|||||||||||||||||||:|||||||:|||
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | ||||||||||||:|||||||||:||||||||||||:|||||||||| |||||||||:
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             |||||||||||:|||:|||:|||||||| :| ||||||||:||||||||||||||||||
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||| ||  ||| |||||| |||||||| || ||||||
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
             |||||||||||||||||  || ||||| ||||||||||||||||
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 90
ACH90651
ID   ACH90651 standard; DNA; 1117 BP.
XX
AC   ACH90651;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #23846.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
```

```
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 23846; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 1117 BP; 367 A; 241 C; 263 G; 246 T; 0 U; 0 Other;

  Query Match            12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity  87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  |||||||||:||||||||   |||||:|||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             ||||||||| :||:|||::|||||| |||||||||:||:|||||||:|||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:||||||: ||| |||||||::||||::|||||||:||||||||:|||||||||||||

Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:|||||||||||: |||||||| ||||||||||||::||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| ||||||||| :|||||||||||||||||||||||||||||||:|||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | ||||||:| ||||||||||:|||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||| ||||||||||:||||:|||||: |||||||||||:|||| ||||||||||||
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:|||||||||||| |||||||
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             |||||||||||:|||||||||||||||||||||||||||||||||||||:||||:|||||||
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   |||||:||||||||:|||||||||||||||
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:||||||:|||||||| ||||||||||::||||||||:|:|||||||| ||
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:||||||||||:|||||||||||||||||||||||:|||||||:|||
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | |||||||||||||:|||||||:||||||||||||:|||||||||| |||||||||:
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|||:|||||||| :| |||||||:|||||||||||||||||||||
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||||||| ||   ||| ||||||  |||||||| || ||||||
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
             |||||||||||||||||  || ||||| ||||||||||||||||
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 91
AAT02815
ID   AAT02815 standard; cDNA; 1158 BP.
XX
AC   AAT02815;
XX
DT   20-MAR-1996  (first entry)
XX
DE   Multiple sclerosis associated retrovirus clone MSRV-1B pol gene.
XX
```

```
KW   Pol gene; multiple sclerosis; infectious agent; PCR; amplification; ss;
KW   retrovirus; ERV-9; HSERV-9; endogenous; MSRV-2; probe; primer; antibody.
XX
OS   Synthetic.
XX
PN   WO9521256-A1.
XX
PD   10-AUG-1995.
XX
PF   06-FEB-1995;   95WO-FR000142.
XX
PR   04-FEB-1994;   94FR-00001529.
PR   04-FEB-1994;   94FR-00001530.
PR   04-FEB-1994;   94FR-00001531.
PR   04-FEB-1994;   94FR-00001532.
PR   24-NOV-1994;   94FR-00014322.
PR   23-DEC-1994;   94FR-00015810.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Mallet F,  Mandrand B,  Bedin F,  Beseme F;
XX
DR   WPI; 1995-283776/37.
XX
PT   Two pathogenic or infectious agents associated with multiple sclerosis -
PT   one being a retrovirus, useful for early diagnosis, prevention and
PT   treatment.
XX
PS   Claim 3; Page 57; 104pp; French.
XX
CC   The nucleotide sequence of the pol gene from the multiple sclerosis
CC   associated infectious agent clone designated MSRV-1. The sequence is a
CC   compilation of the sequences from other MSRV-1 clones (T022816-23). The
CC   MSRV-1 and MSRV-2 clones (T0224-6) are derived from the same strain of
CC   infectious agent, either POL2 or MS7PG or their variants, and can be
CC   produced by the same cell line, LM7PC or PLI-2 or other cell culture
CC   infected with the multiple sclerosis associated agents. The clones were
CC   isolated by extracting nucleic acids from cells of multiple sclerosis
CC   patients, converting the nucleic acid to cDNA and PCR amplifying a region
CC   of the pol gene conserved in retroviruses. The clones isolated fell into
CC   2 distinct types: one was similar to the ERV-9 and HSERV-9 endogenous
CC   retroviruses and designated MSRV-1B; the others were highly homologous to
CC   the known agent MSRV-2 and designated MSRV-2B. The nucleic acid, probes
CC   or primers derived from these sequences can be used for detecting agents
CC   associated with multiple sclerosis. The probes may be used to
CC   discriminate between these agents, opt. in conjunction with amplification
CC   primers and antibodies specific for the infectious agents
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;

  Query Match            12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity  86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  |||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||||| |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             ||||||||||||||||:|||||||||||||||||||||||:||||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db         317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||||| :|
Db         377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db         437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||||
Db         497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db         557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db         617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db         676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  |||||||||:||||||||
Db         736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db         791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:||||||||||||::||| ::||||||:||||
Db         850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db         910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db         970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||| ||||||||||: ||||
Db        1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db        1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db        1150 CTTAAACC 1157


RESULT 92
AAT37973
ID   AAT37973 standard; cDNA; 1158 BP.
XX
AC   AAT37973;
XX
DT   27-AUG-2003  (revised)
DT   07-MAY-1997  (first entry)
XX
DE   Partial sequence of MSRV-1 POL gene.
XX
```

```
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS   Retroviridae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1158
FT                   /*tag=  a
FT                   /product= "fragment of MSRV-1 polymerase"
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
DR   P-PSDB; AAW12370.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 24; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. The retroviruses were
CC   isolated from the cell lines by fractionating the cells and assaying for
CC   reverse transcriptase activity. Messenger RNA was then purified and
CC   amplified by RT-PCR using primers based on reasonably conserved sequences
CC   found in the pol genes of known exogenous and endogenous retroviruses and
CC   other viruses having reverse transcriptase activity. The resultant
CC   amplified cDNAs were cloned and sequenced. Sequence analysis separated
CC   the clones into two groups: the first has homology to the human
CC   endogenous retrovirus HSERV-9 and is designated MSRV-1B (see AAT37974-79)
CC   ; the second group has homology to the retrovirus MSRV-2A (AAT38002)
CC   isolated from LMP7 cells (from sclerotic plaques) and is designated MSRV-
CC   2B. The sequence presented here is a composite of the sequence
CC   corresponding to the pol gene from the MSRV-1 group (AAT37974-81). The
CC   sequences can be used for the diagnosis, prevention and treatment of RPA
CC   caused by infection or reactivation of the retroviruses. (Updated on 27-
CC   AUG-2003 to correct OS field.)
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             ||||  ||||| ||||| | | ||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db         197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             ||||||||||||||||:|||||||||||||||||||||||:||||||||||||||||||
Db         257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             ||||||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db         317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db         377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db         437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:||||||||||||||||:|||||||||||||||
Db         497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||:|||||||||||||||||:|::|
Db         557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db         617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| ||||||||||||||||||||||| |||||:|||||||||:|
Db         676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db         736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db         791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||||::||| ::||||||:||||
Db         850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:|||||  ||| ||||||||||
Db         910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||:|||||||| ||||||||||||||||
Db         970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             ||||:|||||::|||||||||||||||||||||||||||||||||| ||||||||||: ||||
Db        1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db        1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db        1150 CTTAAACC 1157


RESULT 93
AAT96527
ID   AAT96527 standard; cDNA; 1158 BP.
```

```
XX
AC   AAT96527;
XX
DT   11-MAR-1998  (first entry)
XX
DE   MSRV-1B pol* sequence.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; primer; PCR; ss;
KW   amplification; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1158
FT                   /*tag=  a
FT                   /product= "pol_protein"
FT                   /transl_except= (pos: 514. .516, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 661. .663, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 694. .696, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 943. .945, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 958. .960, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 973. .975, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 1015. .1017, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 1095. .1097, aa: Xaa)
FT                   /note= "Xaa = unknown"
FT                   /transl_except= (pos: 1129. .1131, aa: Xaa)
FT                   /note= "Xaa = unknown; no stop codon is given at the end
FT                   of the sequence"
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
DR   P-PSDB; AAW36039.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Claim 2; Fig 9; 188pp; French.
XX
CC   This is the nucleotide sequence of a region designated pol* from the
CC   multiple sclerosis related virus 1B (MSRV-1B). MSRV-1 sequences are
CC   associated with multiple sclerosis and some are also associated with
CC   rheumatoid arthritis. Nucleic acid sequences from the MSRV-1 virus can be
CC   used in the diagnosis of multiple sclerosis and rheumatoid arthritis.
CC   Note: this sequence is specifically excluded from the claimed sequences
CC   of the specification
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                ||||   |  | |  ||||||||  ||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                ||||  |||||  |||||  | |  ||||||||||||||||||  |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||:|||||||||||:|:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||||  ||  | | |||  |||||||:|||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
                | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:|||||||||||| |||:||  |||||:|||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||:|:|||| |||||:|||||  ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:||||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
```

```
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                |||||  ||
Db       1150 CTTAAACC 1157


RESULT 94
AAV43115
ID   AAV43115 standard; cDNA; 1158 BP.
XX
AC   AAV43115;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus-1 pol gene sequence.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1158
FT                   /*tag=  a
FT                   /product= "MSRV-1 pol protein"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71049.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 5; Fig 9; 286pp; English.
XX
CC   The present sequence is representative of the multiple sclerosis (MS)
CC   associated retrovirus-1 (MSRV-1) pol gene sequence. The invention
CC   provides complete or partial genomic sequences of the MSRV-1 pol gene,
CC   gag gene and env gene, and polypeptides encoded by these genes. The
CC   invention also provides antibodies raised against the polypeptides. The
CC   genomic sequences, polypeptides and antibodies are also claimed useful
CC   for diagnosing infection by MS and rheumatoid arthritis-associated
CC   viruses, and also for prevention and treatment of infection with these
CC   viruses
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  ||||||||  |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::||||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :|||||:|||||:|||||| |||||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||  ||  | | |||  ||||||||:|||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
                | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:||||||||||| |||:|| |||||:|||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                |||||  ||
Db       1150 CTTAAACC 1157
```

```
RESULT 95
ADB84549
ID   ADB84549 standard; DNA; 1158 BP.
XX
AC   ADB84549;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol gene consensus sequence #1.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84519.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 5; Fig 9; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;
```

```
  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity  86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   | | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:||||||||||:|||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||| |||||||||||||||||||||||||||||||| ||| |||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             ||||||||||||||||:|||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:|||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
```

```
                |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                ||||| |:||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                ||||| ||
Db       1150 CTTAAACC 1157


RESULT 96
ADG14742
ID   ADG14742 standard; cDNA; 1158 BP.
XX
AC   ADG14742;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1 pol gene.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 5; SEQ ID NO 1; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 1158 BP; 324 A; 286 C; 241 G; 298 T; 0 U; 9 Other;

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||  |  | |  ||||||||  |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                ||||||||| |||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :|||||:|||||:|||||| |||||||||||||||||||| ||||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||||  ||  | | |||  |||||||||:|||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
                | ||:||||||||:|||||||| ||||:||:|||:: ||||| ||||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:|||||||||||| |||:||  |||||:||||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||||||:|:|||| |||||:|||||  ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                ||||| |:||||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                ||||| ||
Db       1150 CTTAAACC 1157


RESULT 97
ABN97933/c
ID   ABN97933 standard; DNA; 1216 BP.
XX
```

```
AC   ABN97933;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Human retroviral sequence HE4.
XX
KW   Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW   multiple sclerosis; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO9967395-A1.
XX
PD   29-DEC-1999.
XX
PF   23-JUN-1999;   99WO-FR001513.
XX
PR   23-JUN-1998;   98FR-00007920.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Alliel PM,  Perin J,  Rieger F;
XX
DR   WPI; 2000-160587/14.
XX
PT   New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT   for diagnosis, treatment and prevention of autoimmune and neurological
PT   diseases.
XX
PS   Claim 3; Page 135; 225pp; French.
XX
CC   The present invention relates to new nucleic acid sequences of human
CC   endogenouse retrovirus, HERV-7q, which is located on chromosome 7q.
CC   Regulatory elements associated with HERV-7q may alter expression of other
CC   genes (even remote genes) on the same chromosome, inducing immunological
CC   and/or neurological changes (which may be pathological or protective/
CC   curative). HERV-7q peptides can be used to improve efficiency of the
CC   immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC   sequences can be used in immunogenic or vaccinating compositions, for
CC   protection against autoimmune diseases, particularly multiple sclerosis.
CC   The peptides may also be used (by sequence comparison) to detect/identify
CC   endogenous retroviruses that are abnormally expressed in cancer,
CC   neuropathologies or other autoimmune diseases. The present sequence was
CC   used to illustrate the invention
XX
SQ   Sequence 1216 BP; 315 A; 221 C; 331 G; 349 T; 0 U; 0 Other;

  Query Match             11.9%;  Score 894.401;  DB 1;  Length 1216;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches 1003;  Conservative   21;  Mismatches  102;  Indels   60;  Gaps    3;

Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATT 4729
             ||||||||||||||:||||||:|||||||||||||| | |:|:||:|||||||||| ||||
Db       1216 AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACTCAGGAGAAGACTGGCAACTAGATT 1157

Qy       4730 TTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCA 4789
             |||||||| ||||||| ||||||||||||||||||||||||||||||:||| |||||
Db       1156 TTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATATTTTCA 1097

Qy       4790 CGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCAC 4849
             | |||||||||:||||||||  ||||||||||||||| |||||| ||||||||||||||
Db       1096 CTGGTTGGGCAGAGGCCTTCCTTTGTAGGACAGAAAAGGACCAAGATGTAATAAAGGCAC 1037

Qy       4850 TAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCC 4909
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1036 TAATTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATGGCC 977

Qy       4910 CTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACT 4969
             | | |||| |||| |||||||||||||||||||||||||| ||||| |||| ||||||||
Db        976 CCGTTTTCAAGGCTACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCACT 917

Qy       4970 TACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCA 5029
             |||||||||| || |||||||| |||||||| || ||| ||||||||||||| :|||||

Db        916 TACACTGCGCGTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAATGAAGCACTCA 857

Qy       5030 AAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTA 5089
             || ||||| ||||||||| |||||| |||||||||||  |||||||||:|||||||||||
Db        856 AACGACATTTAAAAAAGCTAACCCAAGAAACCCACCTTGCATGGCCTGCTCTGTTGCCTA 797

Qy       5090 TAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGC 5149
             |||||||  |||||||   |||| ||||||||||||||||||||||||||||| ||||||
Db        796 CAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAAGCAGGACTTAGCCCATACAAAATGC 737

Qy       5150 TGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAG 5208
             |||||||| |||||||| ||||||| |||||||||||||||| | |||  |||| |||||
Db        736 TGTATGGATGGCCCTTCCTAACCAAGGACCTTGTGCTTGACCGAGAGATGGCCAGCTTAG 677

Qy       5209 TTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCT 5268
             |||||||||||||||||| ||||||||||||||||||||||||||||||| |||||||
Db        676 TTGCAGACATCACCTCCTTGGCCAAATATCAACAAGTTCTTAAAACATTACAGGGAACCT 617

Qy       5269 ATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCC 5328
             | || ||||||||||||||  || |||| |||||: ||||||||||||||||||||||||
Db        616 GTTCCCGAGAAGAGGGAAAGTAATTATTGCACCCTGGTGACATGGTATTAGTCAAGTCCC 557

Qy       5329 TTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 5388
             |||: ||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db        556 TTCCTTCTAAGTCCCCATCCTTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 497

Qy       5389 YTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 5448
             :|||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        496 CTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGAATACATCACACTCGAGTCA 437

Qy       5449 AATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 5508
             ||  ||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        436 AACTCTGTATACTGCCAAAGGAACCCGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 377

Qy       5509 AACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAAT 5568
             |||||||||||||| ||| |||||||||||| | ||||||  |||||||||||||| |||
Db        376 AACCTCTAGAGGATCTGCACCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAAT 317

Qy       5569 CATAAATCCCCCATG---------------------------------------------- 5583
             | || |||| |   |  .
Db        316 CGTAGATCCTCATGGCCCTCCTTGTCATATTTTTTTCTTTACTGTTCTCTTACCCCCTTT 257

Qy       5584 ------------GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTT 5630
                         :|||||||||||||||||||||||||:||:||||:|||:|||||:|||
Db        256 CACTCTCACTGAACCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTT 197

Qy       5631 CACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTT 5690
             |||||||||||||||||||||||||| |||||  |||||||||||| |:|||:|||| |
Db        196 CACTCTCACTGCACCCCCTCCATGCCACTGTACTACCAGTAGCTCCCATTACCAAGAGCT 137

Qy       5691 TCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAA 5750
             |||||||||||||| || ||| |||||||||||||| |||||||||||||||| |||:|||
Db        136 TCTATGGAGAATGCGGCTTCCTGGAAATATTGATGACCCATCGTATAGGAGTTTTTCTAA 77

Qy       5751 GGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||| |||||  |||||  |||||||| ||||| |||   ||| |
Db         76 AGGAAACCCCATTTTCACCACCCACACCTATATGACCCTGCACTTC 31


RESULT 98
AAS71725
ID   AAS71725 standard; cDNA; 910 BP.
XX
AC   AAS71725;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #7529.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG07538.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 7529; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 910 BP; 255 A; 243 C; 180 G; 232 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 714.8;  DB 1;  Length 910;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  821;  Conservative   25;  Mismatches   52;  Indels   40;  Gaps    7;

Qy       2238 TTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 2297
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 TTCTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 65

Qy       2298 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 2357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 125

Qy       2358 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAA 2417
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        126 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAA 184

Qy       2418 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCT 2477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCT 244

Qy       2478 GTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGA 2537
             |||||||||||||||||||||  | |||||||||||||||||||| |||||||| |||
Db        245 GTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGCAATACCCCCTTAGA 304

Qy       2538 CCCAAGGCCCAACAA--GGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||  ||||||||||||  || ||||||||||||||||||||| ||||||||||||||| |||
Db        305 CCTGAGGCCCAACAAAGGGCTTCCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 364

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             ||||||||||||  |  ||||||||||| | |        |||    |  | ||||||||||
Db        365 TAAAACCATGCAATAGCCCCTGCAGTACTCC-------AATTTTAGGAGTACAGAAACCC 417

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| ||||||||||||||||||||||||||| || |||||| |||||| |
Db        418 AACGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCGAT-GAGGCCATTGTCCCTC 475

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||  |||||||||| ||::|||||||||||||||||||||||||||
Db        476 TATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGT 535

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             |||||||:|||||||||||:|||||||||| ||||||||||||||||||||||| ||||
Db        536 TGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGAGTCTC 595

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| |||||| ||||:  |||||||||||||||||:|||
Db        596 AATTCTTGTTTGCCTTTGAAGATCCTTCAATCCCAATGTCTCAACTCACCTGGACTGTTT 655

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             ||||||||||||||||||||:|||||||||||  |||  ||||||||||||||||||||
Db        656 TACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTCACCATACATTAGCCCAAGACTTGA 715

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |:||:|::|||||||||||||||||||||||||||||||||||||||||| || ||:
Db        716 GCCAGTTCTCATACCTGGACACTCTTGTCC----------------------TTTGCCT 752

Qy       3016 CCYRTTCAGAAACC-----TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGC 3070
             ||::|| |||||||       |||||||||||||||||||||| :|||||:||| |||||||
Db        753 CCCGTTGAGAAACCAACAGGAGTGCCATCAAGCCACCCAAGTGCTCTTAAATGTCCTCGC 812

Qy       3071 YACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATA 3130
             :|||||||||||||:|||||||||||| |:|:|||||:|||||||||||||||||||||||
Db        813 TACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATA 872

Qy       3131 CTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
             ||||||: |||:|||||||||||| ||||:||||||||
Db        873 CTTAGGGGTAAAATTATCCAAAGGGACCAGGGCCCTCA 910


RESULT 99
AAI23803
ID   AAI23803 standard; DNA; 792 BP.
XX
AC   AAI23803;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13736 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
```

```
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 13736; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             |||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||||||| |||||||||||  ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 100
ABA68919
ID   ABA68919 standard; DNA; 792 BP.
XX
AC   ABA68919;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #17224.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 17224; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120
```

```
Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             ||||||||||||||||||||||||||||||||||| |||||||||| ||||    ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
              |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| |||||||| ||||||| ||   ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792
```

```
RESULT 101
AAI49115
ID   AAI49115 standard; DNA; 792 BP.
XX
AC   AAI49115;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17801 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 17801; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;
```

```
  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| ||||||||  ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             ||||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| |||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             ||||||||||||||||||||||||||||||||||| |||||||||| ||||    ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
              |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657
```

```
Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             ||||||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 102
ABA50940
ID   ABA50940 standard; DNA; 792 BP.
XX
AC   ABA50940;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9635.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 9635; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX

SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||  |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             |||||||||  |||||||||||| |||| |:|||||  |||||||||| |||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| |||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
              |||||||||||||||| | | ||| |||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             |||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 103
ABA35872
ID   ABA35872 standard; DNA; 792 BP.
XX
AC   ABA35872;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #14338 for gene expression analysis in human heart cell sample.
XX
```

```
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 14338; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| |||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 104
AAK43040
ID   AAK43040 standard; DNA; 792 BP.
XX
AC   AAK43040;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17597.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 17597; 658pp + Sequence Listing; English.
XX
```

```
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match            9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 105
AAK17251
ID   AAK17251 standard; DNA; 792 BP.
XX
AC   AAK17251;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 17242.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 17242; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match            9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
```

```
                ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
                |||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
                |||| |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
                ||||| | |||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
                  |||||||||||||||| | | ||| |||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
                |||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
                ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
                ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 106
ABS42675
ID   ABS42675 standard; DNA; 792 BP.
XX
AC   ABS42675;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 17665.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 17665; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
                |||||||| |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
                |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
                |||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
                |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
                ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
                |||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
                |||| |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
                ||||| | |||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
                  |||||||||||||||| | | ||| |||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
                |||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717
```

```
Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 107
AAI09412
ID   AAI09412 standard; DNA; 792 BP.
XX
AC   AAI09412;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #9403 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 9403; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             ||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||| |||||||||| |||| ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||||| |||||||||||||||| |||||||| |||||| || ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 108
ABS17105
ID   ABS17105 standard; DNA; 792 BP.
XX
AC   ABS17105;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 17096.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 17096; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 792 BP; 228 A; 223 C; 134 G; 207 T; 0 U; 0 Other;

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
              |||||||||  |||| |||||||||||||||||| ||| |||||||| ||||| || ||||

Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
              ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||  ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
              |||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
              |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
              ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
              |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
              |||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
              |||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 109
AAI18958
ID   AAI18958 standard; DNA; 944 BP.
XX
AC   AAI18958;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #8891 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
```

```
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 8891; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 944 BP; 260 A; 261 C; 195 G; 228 T; 0 U; 0 Other;

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || ||||||||||||||| |||| ||||||| ||
Db             1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db            61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db            121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||||||| | |    ||| ||||||||||||  |
Db            181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||| |||| |||||||||||| |
Db            241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:||||||||||||||| || ||
Db            300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||||| |||||||||||| |||||
Db            356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
```

```
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
                |||||||||||||||||||||||||||||| |||||| |||||||||| || || | |
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
             | |   ||||||| |||||||||||||||||||||||||| ||  | ||||| || ||
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
             | |         |||    |  | |||||||||| |||||||||| ||||||||||||
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             |||||||||||||| | |||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||| |||||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||| |||||| |||||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
             |||
Db        941 CGAA 944
```

```
RESULT 110
AAI44090
ID   AAI44090 standard; DNA; 944 BP.
XX
AC   AAI44090;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #12776 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
```

```
PS   Claim 25; SEQ ID NO 12776; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 944 BP; 260 A; 261 C; 195 G; 228 T; 0 U; 0 Other;

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || |||||||||||||| |||| ||||||| ||
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             |||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||| | |   ||| ||||||||||||   |
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||||| |||| ||||||||||| |
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:|||||||||||||||| || ||
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||||| |||||||||||| |||||
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
                |||||||||||||||||||||||||||||| |||||| |||||||||| || || | |
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
             | |   ||||||  |||||||||||||||||||||||||||  ||  | ||||| || ||
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
               | |        |||   |  | |||||||||| |||||||||| ||||||||||||
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             |||||||||||||| | |||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||||| |||||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||||||| |||||| |||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
               | ||
Db        941 CGAA 944


RESULT 111
ABS37794
ID   ABS37794 standard; DNA; 944 BP.
XX
AC   ABS37794;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 12784.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 12784; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 944 BP; 260 A; 261 C; 195 G; 228 T; 0 U; 0 Other;

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
```

```
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025

Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085

Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142

Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202

Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262

Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322

Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382

Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442

Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502

Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562

Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622

Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682

Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742

Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802

Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862

Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866

Db        941 CGAA 944


RESULT 112
ABS12201
ID   ABS12201 standard; DNA; 944 BP.
XX
AC   ABS12201;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 12192.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 12192; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
```

```
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 944 BP; 260 A; 261 C; 195 G; 228 T; 0 U; 0 Other;

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || |||||||||||||| |||| |||||||  ||
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             |||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||| | |   ||| ||||||||||||   |
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||| |||| ||||||||||| |
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:|||||||||||||||| || ||
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||||| |||||||||||| |||||
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
                |||||||||||||||||||||||||||||| |||||| |||||||||| || || | |
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
             | |   ||||||| ||||||||||||||||||||||||||| ||  | ||||| || ||
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
               | |       |||   |  | |||||||||| |||||||||| ||||||||||||||
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             |||||||||||||| | |||||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||||| |||||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||||||| |||||| |||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
             | ||
Db        941 CGAA 944


RESULT 113
AAS72224
ID   AAS72224 standard; cDNA; 924 BP.
XX
AC   AAS72224;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8028.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08037.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8028; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 924 BP; 272 A; 248 C; 179 G; 225 T; 0 U; 0 Other;

  Query Match             8.8%;  Score 662.8;  DB 1;  Length 924;
```

```
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  724;  Conservative   44;  Mismatches   71;  Indels   10;  Gaps    4;

Qy       2494 AGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAG 2553
             ||||   |  |   | |    |  |||||||||||||| ||||||||||||||||||||
Db         86 AGGACCACTTTCTTTGTGGTCAAGAAAGGCAGTACCCCCTTAGACCCAAGGCCCAACAAG 145

Qy       2554 GATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTC 2613
             ||  |||||||  | ||||||||||| |||||||||||||| ||||||||||||||   |  |
Db        146 GACTCCAAAAAACTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCC 205

Qy       2614 CCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTT 2673
             ||||||  ||  | |          | ||   |  | ||||||||||  |||||||||| ||||
Db        206 CCTGCAATACTCC-------AACTGTAGGAGTACAGAAACCCAACGGACAGTGGA-GGTT 257

Qy       2674 AGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTA 2733
             |||||||  ||||||||||||||||||  |||  |  |||||  |  |  ||||||||||| |||||
Db        258 AGTGCAAAATCTCAGGATTATCAAT-GAGACTGTTGTTCCTCTATACCCAGCTGAACCTA 316

Qy       2734 GCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACC 2793
             ||||||||||  ||::|||||||||||||||||||||||||||||||  ||  |:||||||||||
Db        317 ACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTATATAGTCCTGGACC 376

Qy       2794 TTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTG 2853
             ||:|||||||||||||||||||||||||||||||||||||  ||  ||||||  ||||||||||||
Db        377 TTAAGGATGCCTTCTTCTGCATCCCTGTACATCCTGATTCCCAATTCCTGTTTGCCTTTG 436

Qy       2854 AAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||  ||||||||  ||:|  ||||||||||||||||||||:||||||  ||||||||||||||
Db        437 AAGATCCTTCAAACTCAACGTCTCAACTCACCTGGACTGTTTTACGCCAAGGGTTCAGGG 496

Qy       2914 ATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGG 2973
             ||||:|||||||||||||||||||||||||||||||||  |||||:||:|::||||||||||
Db        497 ATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGATTTGAGCCAATTCTCATACCTGG 556

Qy       2974 ACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGT 3033
             ||||||||||||||  :|||:|||||||||||||||||:||:  ||::|||||||||||||||
Db        557 ACACTCTTGTCCTTTGGTATGTGGATGATTTACTTTTAGCCACCTGTTCAGAAACCTTGT 616

Qy       3034 GCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCA 3093
             ||||||||||||||||  :|||||:|||||||||  |:|||||||||||||:||||||||
Db        617 GCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCA 676

Qy       3094 AACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAG 3153
             |||:|:|:|||||:||||||||||||||||||||||||||||||:||||:||||||||||||
Db        677 AACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAG 736

Qy       3154 GCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAA 3213
             ||   |:||||||||||||||||:::|||||||||||||||||||||||||||||| ||:|||||
Db        737 GCTTGAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCTTCCCAAAA 796

Qy       3214 CCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCC 3273
             |||||||||||||||:|::|||||||||:|||:|||:|||| |||||: |||||| |||
Db        797 CCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTACCGAATGTGGATT-CCC 855

Qy       3274 AGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAAT 3333
             ||||:  ||:|||||| ||| :|||||:|||||:|||||| ||||||||||||||||||
Db        856 AGGTACAGCAAAATAGGCAGATCATTATATACACTAATTAAAGAAACTCAGAAAGCCAAT 915

Qy       3334 ACCCATTTA 3342
             |||||||||
Db        916 ACCCATTTA 924


RESULT 114
AAS71722/c
ID   AAS71722 standard; cDNA; 902 BP.
XX
AC   AAS71722;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #7526.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG07535.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 7526; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 902 BP; 232 A; 169 C; 228 G; 273 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 646;  DB 1;  Length 902;
  Best Local Similarity   93.6%;  Pred. No. 0;
  Matches  661;  Conservative   10;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       6165 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 6224
             ||||  || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 GCCACATATTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 647

Qy       6225 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCT 6284
             |||||||  |||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        646 CACTTCCATTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCT 587

Qy       6285 CACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 6344
             ||||  |||||||||||||||||||| | |||||||||||||||  |||||||| |||||
Db        586 CACCCGTGTAAAATTTAGCAATACTATAGACACAACCAACTCCCAGTGCATCAGATGGGT 527

Qy       6345 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 6404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 467
```

```
Qy       6405 AGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCC 6464
             ||||||||  ||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db        466 AGCCTATCACTGTTTGAATGGCTTTTCGGAATCTATGTGCTTCCTCTCATTCTTAGTGCA 407

Qy       6465 CCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
             |||:|||:|||||||||||||||||||||||||||| ||||||| || ||||||||||||
Db        406 CCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTACCTAAGCCCCGCAA 347

Qy       6525 CAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTAC 6584
             |||||||||||||||||||||||||||||  ||||||| |||||||   | |||||||
Db        346 CAAAAGAGTACCCATTCTTCCTTTTGTTATCCGAGCAGGAATGCTAGGCAGATTAGGTAC 287

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             |||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGCATGGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 227

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
             ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        226 TGGTGACATGGAACGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 167

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGG 6764
             |||||||||||| ||||:|||||||||||||||||||:|||||| || ||:||||||||
Db        166 AGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCAGAAAGAGGGGG 107

Qy       6765 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 6824
             ||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db        106 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCAC 47

Qy       6825 TGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGAR 6870
             ||||||||||:|||||||||:|||| |||||||||| |:|||| |:
Db         46 TGAGAAAGTTAAAGAAATTCGAGATGGAATACAACGCAGAGCATAG 1
```

```
RESULT 115
AAS76467
ID   AAS76467 standard; cDNA; 841 BP.
XX
AC   AAS76467;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12271.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12280.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 12271; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 841 BP; 241 A; 224 C; 169 G; 207 T; 0 U; 0 Other;

  Query Match             8.5%;  Score 637.7;  DB 1;  Length 841;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  750;  Conservative   24;  Mismatches   62;  Indels   39;  Gaps    7;

Qy       2299 ATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATT 2358
             |||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATT 60

Qy       2359 ATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAAT 2418
             |||||||||||||||||||||||||||||||| ||||||||| |||||||||| |||||||||
Db         61 ATACACCTGAACATAGGAGAAGGAACACCTGTTTGTTGT-CCCCTGCTTAAGGAAGGAAT 119

Qy       2419 TAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTG 2478
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||
Db         120 TAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCCATCCTG 178

Qy       2479 TTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGAC 2538
             |||||||||||||||||||||  | |||||||||||||||||| |||||||| ||||
Db         179 TTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGCAATACCCCCTTAGAC 238

Qy       2539 CCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAA 2598
             |  |||||||||||||| |||||||||||||||||||||| ||||||  |||||| ||||||
Db         239 CTGAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTAAAAGTGCAAGGCCTAGTAA 298

Qy       2599 AACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGT 2658
             || ||||||  |  ||||||  || | |        |||   |  | ||||||||||| |
Db         299 AAACATGCAATAGCCCCTGCGATACTCC-------AATTTTAGGAGTACAGAAACCCAAT 351

Qy       2659 GGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTAT 2718
             |||||||||| |||| |||||||||||||||||| || |||||| |||||| | |||
Db         352 GGACAGTGGA-GGTTGGTGCAAGATCTCAGGATTATCGAT-GAGGCCATTGTCCCTCTAT 409

Qy       2719 ACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGT 2778
             |||||||||||||| |||||||||| ||::|||||||||||||||||||||||||||| ||
Db         410 ACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTTGT 469

Qy       2779 TTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAAT 2838
             |||||:|||||||||||:|||||||||| |||||||||||||||||||||| |||||||
Db         470 TTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGAGTCTCAAT 529

Qy       2839 TCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTAC 2898
             |||||||||||||||||||| |||||| ||||:  |||||||||||||||||:||||||
Db         530 TCTTGTTTGCCTTTGAAGATCCTTCAATCCCAATGTCTCAACTCACCTGGACTGTTTTAC 589

Qy       2899 CCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYC 2958
             ||||||||||||||||||:|||||||||||  |||  |||||||||||||||||||:|
Db         590 CCCAAGGGTTCAGGGATAGTCCCCATCTATTTCACCATACATTAGCCCAAGACTTGAGCC 649

Qy       2959 ARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCY 3018
             |:|::|||||||||||||||||||||||                  || ||: ||:
Db         650 AGTTCTCATACCTGGACACTCTTGTCC----------------------TTTGCCTCCC 686
```

```
Qy       3019 RTTCAGAAACC-----TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
              :|| |||||||       |||||||||||||||||||| :|||||:||| |||||||:|| 
Db        687 GTTGAGAAACCAACAGGAGTGCCATCAAGCCACCCAAGTGCTCTTAAATGTCCTCGCTAC 746

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTT 3133
              |||||||||||:||||||||||| |:|:|||||:|||||||||||||||||||||||||| 
Db        747 CTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTT 806

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
              |||: |||:|||||||||||| ||||:|||||||| 
Db        807 AGGGGTAAAATTATCCAAAGGGACCAGGGCCCTCA 841


RESULT 116
AAT02823
ID   AAT02823 standard; cDNA; 741 BP.
XX
AC   AAT02823;
XX
DT   20-MAR-1996  (first entry)
XX
DE   Multiple sclerosis associated retrovirus clone MSRV-1B PSJ-17.
XX
KW   Pol gene; multiple sclerosis; infectious agent; PCR; amplification; ss;
KW   retrovirus; ERV-9; HSERV-9; endogenous; MSRV-2; probe; primer; antibody.
XX
OS   Synthetic.
XX
PN   WO9521256-A1.
XX
PD   10-AUG-1995.
XX
PF   06-FEB-1995;   95WO-FR000142.
XX
PR   04-FEB-1994;   94FR-00001529.
PR   04-FEB-1994;   94FR-00001530.
PR   04-FEB-1994;   94FR-00001531.
PR   04-FEB-1994;   94FR-00001532.
PR   24-NOV-1994;   94FR-00014322.
PR   23-DEC-1994;   94FR-00015810.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Mallet F,  Mandrand B,  Bedin F,  Beseme F;
XX
DR   WPI; 1995-283776/37.
XX
PT   Two pathogenic or infectious agents associated with multiple sclerosis -
PT   one being a retrovirus, useful for early diagnosis, prevention and
PT   treatment.
XX
PS   Claim 3; Page 59-60; 104pp; French.
XX
CC   The partial nucleotide sequence of the pol gene from the multiple
CC   sclerosis associated infectious agent clone designated MSRV-1B PSJ-17.
CC   The MSRV-1B (AAT02815-23) and MSRV-2 clones (AAT02824-6) are derived from
CC   the same strain of infectious agent, either POL2 or MS7PG or their
CC   variants, and can be produced by the same cell line, LM7PC or PLI-2 or
CC   other cell culture infected with the multiple sclerosis associated
CC   agents. The clones were isolated by extracting nucleic acids from cells
CC   of multiple sclerosis patients, converting the nucleic acid to cDNA and
CC   PCR amplifying a region of the pol gene conserved in retroviruses. The
CC   clones isolated fell into 2 distinct types: one was similar to the ERV-9
CC   and HSERV-9 endogenous retroviruses and designated MSRV-1B; the others
CC   were highly homologous to the known agent MSRV-2 and designated MSRV-2B.
CC   The nucleic acid, probes or primers derived from these sequences can be
CC   used for detecting agents associated with multiple sclerosis. The probes
CC   may be used to discriminate between these agents, opt. in conjunction
CC   with amplification primers and antibodies specific for the infectious
CC   agents
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;


  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
              ||||||||||||| :|||||:||||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:|||||||||:|      | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||||| ||:||||||||:||||||| ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:||||||||||||||||:|||:|||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 117
AAT37981
ID   AAT37981 standard; cDNA; 741 BP.
XX
AC   AAT37981;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   Rheumatoid polyarthritis associated retrovirus pol gene clone PSJ17.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
```

```
XX
OS   Retroviridae.
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 27; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is clone clone
CC   PSJ17 which corresponds a fragment of the pol gene from the viruses. The
CC   clone was isolated using the endogenous reverse transcriptase activity of
CC   virions isolated from PLI-2 cells to generate cDNA from mRNA purified
CC   from PLI-2 cells. The viral sequences can be used for the diagnosis,
CC   prevention and treatment of RPA caused by infection or reactivation of
CC   the retroviruses. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
               |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:||||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:||||||:||||||||||||: 

Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: ||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 118
AAT96517
ID   AAT96517 standard; cDNA; 741 BP.
XX
AC   AAT96517;
XX
DT   06-MAR-1998  (first entry)
XX
DE   Clone PSJ17 comprising a MSRV-1 type retrovirus.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; primer; PCR; ss;
KW   amplification; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Example 3; Page 96; 188pp; French.
XX
CC   This is the sequence represents a fragment of the genome from the
CC   multiple sclerosis related virus 1B (MSRV-1B), isolated using an
CC   endogenous reverse transcriptase assay to isolate endogenous
CC   retroviruses. The sequence was isolated from the cell line PLI-2 derived
CC   from patients with multiple sclerosis. Sequence analysis shows that this
CC   sequence has homology to certain known retroviral elements, especially
CC   the endogenous human ERV-9 retrovirus. MSRV-1 sequences are associated
CC   with multiple sclerosis and some are also associated with rheumatoid
CC   arthritis. Nucleic acid sequences from the MSRV-1 virus can be used in
CC   the diagnosis of multiple sclerosis and rheumatoid arthritis
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
```

```
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||||| || ||| |||||||||: |||:|||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:|||||||:|||||| ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:||||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: |||||||| |:|||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 119
AAV43123
ID   AAV43123 standard; cDNA; 741 BP.
XX
AC   AAV43123;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus-1 clone PSJ17 DNA sequence.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 3; Fig 6; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) clone PSJ17 DNA sequence. The invention provides
CC   complete or partial genomic sequences of the MSRV-1 pol gene, gag gene
CC   and env gene, and polypeptides encoded by these genes. The invention also
CC   provides antibodies raised against the polypeptides. The genomic
CC   sequences, polypeptides and antibodies are also claimed useful for
CC   diagnosing infection by MS and rheumatoid arthritis-associated viruses,
CC   and also for prevention and treatment of infection with these viruses
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||||| || ||| |||||||||: |||:|||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:|||||||:|||||| ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:||||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594
```

```
Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| ||||||||||||||||||:|||||::||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 120
ADB84557
ID   ADB84557 standard; cDNA; 741 BP.
XX
AC   ADB84557;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1DNA clone PSJ17.
XX
KW   MSRV; ss; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 3; Fig 6; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV cDNA clone. Note: The SEQ ID numbers for the
CC   sequences as displayed in the main body of the patent do not match the
CC   SEQ ID numbers in the sequence listing. Consequently those sequences
CC   mentioned in the claims may not be the sequences the authors intended to
CC   claim.
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||  :|||||:|||||||||  |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:||||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:||| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| ||||||||||||||||||:|||||::||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 121
ADG14750
ID   ADG14750 standard; cDNA; 741 BP.
XX
```

```
AC   ADG14750;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1 clone PSJ17.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 3; SEQ ID NO 9; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 741 BP; 232 A; 170 C; 159 G; 174 T; 0 U; 6 Other;

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||||:|||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
             :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:|||||||:|||||||| |||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:||||||||||||||||:|||:|||||:||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| ||||||||||||||||||:||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: ||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 122
AAV43222
ID   AAV43222 standard; cDNA; 758 BP.
XX
AC   AAV43222;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus fragment 7.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 190; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) genomic fragment used in the method of the invention.
CC   The invention provides complete or partial genomic sequences of the MSRV-
CC   1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC   genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 758 BP; 243 A; 190 C; 155 G; 170 T; 0 U; 0 Other;

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
```

```
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
              ||||| |||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
              |||||||||||:||    |||||| |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
              ||||| | ||||||||||:||||||:|||||||||||||| |: :||:|||||||||| 
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
              |||||||||||| | |||||  |||||||||||| |||||||||||||| |:|||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
              ||||| |||||||  :|| | |  || |||||||||||||| ||||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
              ||||||    |||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| ||||  |||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
              |||||||||| ||||| |||||||| |||||  | || ||| ||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
              |||||||   |||||||||||| |||||| ||||||||  |  |||| ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
              |||||| |||||  |||||||   |||| | ||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
              |||||  |||||| |||| ||| |||||||||||||||||||||||  | | |  |||||
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
              ||||| |||| |||| |||||||| |||||||||||||: ||||||||
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 123
AAX29705
ID   AAX29705 standard; DNA; 758 BP.
XX
AC   AAX29705;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Clone LB16 from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis; ss.
XX
OS   Viruses.
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   P-PSDB; AAW99555.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 1; Fig 11; 83pp; French.
XX
CC   This sequence represents clone LB16 from a novel multiple sclerosis
CC   related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC   prophylactic or therapeutic compositions to inhibit expression of a
CC   multiple sclerosis related virus and/or virus associated with rheumatoid
CC   polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC   17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 758 BP; 243 A; 190 C; 155 G; 170 T; 0 U; 0 Other;

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
              ||||| |||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
              |||||||||||:||    |||||| |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
              ||||| | ||||||||||:||||||:|||||||||||||| |: :||:|||||||||| 
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
              ||||||||||||| | |||||  |||||||||||||| |||||||||||||||| |:|||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
              ||||| |||||||  :|| | |  || |||||||||||||| ||||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
              ||||||    |||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| ||||  |||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
              |||||||||| ||||| |||||||| |||||  | || ||| ||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
              |||||||   |||||||||||| |||||| ||||||||  |  |||| ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
              |||||| |||||  |||||||   |||| | ||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
              |||||  |||||| |||| ||| |||||||||||||||||||||||  | | |  |||||
```

```
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
             ||||| |||| |||| |||||||| |||||||||||||: ||||||||
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 124
ADG14853
ID   ADG14853 standard; cDNA; 758 BP.
XX
AC   ADG14853;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #14.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 112; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 758 BP; 243 A; 190 C; 155 G; 170 T; 0 U; 0 Other;

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             |||||  |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:||    |||||| |||||:  |||||:|:||||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
             |||||  | ||||||||||:||||||:|||||||||||||||| |: :||:|||||||||||
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
             ||||||||||||| | |||||  |||||||||||| |||||||||||||||| |:|||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
             |||||  |||||||  :|| | |  || |||||||||||||||| ||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
             ||||||    |||||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| |||| |||||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
             |||||||||||| ||||| |||||||| |||||  | || ||| |||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
             |||||||   |||||||||||| |||||| |||||||||  |  |||| ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
             ||||||  |||||  |||||||   |||| | ||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
             |||||  ||||||  |||| ||| |||||||||||||||||||||||  | | |  |||||
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
             |||||  |||| |||| |||||||| |||||||||||||||: ||||||||
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 125
AAH04815
ID   AAH04815 standard; cDNA; 745 BP.
XX
AC   AAH04815;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:1650.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
```

```
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 1650; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 745 BP; 214 A; 196 C; 156 G; 176 T; 0 U; 3 Other;

  Query Match             7.5%;  Score 560.3;  DB 1;  Length 745;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  635;  Conservative   62;  Mismatches   47;  Indels   11;  Gaps    8;

Qy       2842 TGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCC 2901
             |||||||||||||||| ||||||||||:  |||||| ||||||||||:|||||||||
Db          1 TGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAATTCACCTGGACTGTTTTACCCC 60

Qy       2902 AAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCART 2961
             |||||||||||||| ||:||||||||||||||||||||||||||||||||||||:||:|
Db         61 AAGGGTTCAGGGACAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGT 120

Qy       2962 YMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTT 3021
             ::||| |||||||||||||||||||||:|||  ||||||||||||||||:||:|||: ||
Db        121 TCTCACACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCTGCCCTTT 180

Qy       3022 CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCT 3081
             ||||||||||||||||||||||||||||||:|||||:|||||||| ||:|||||||||||
Db        181 CAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTTGCCACCTGTGGCT 240

Qy       3082 ACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAA 3141
             || :||||||||||:|:| |||||:|||||||||||||||| ||||||||||||:||||
Db        241 ACCAGGTTTCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAA 300

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
             :||||||| |||||||||:|||||||||||||||||:::| |||||| | ||||||||||
Db        301 AATTATCCTAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATC 360

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAA 3261
             ||||||:| |||||||||||||||||:|::|||||| ||:|||:|||:|||||||| ||
Db        361 CTCATCCC-AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATAACAGGCTTCTGCCAAA 419

Qy       3262 WATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACT 3321
             :||||||| |||||||:||| :||||||||||:|||||:|||||:|||||||||||||||
Db        420 TATGGATT-CCCAGGTATGGTGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACT 478

Qy       3322 CAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCC 3381
             ||||||| |||||||||||||:|||| |||| |:||||||: :|||  ||||||||||||
Db        479 CAGAAAG-CAATACCCATTTAGTAAGGTGGAAACCTGAAGCGGAAGCAGCTTTCCAGGCC 537

Qy       3382 CCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTT 3441
             |  ||  | | |||  ||||||||:||||||||||||:||||||||:||||||||||
Db        538 CTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT-- 593

Qy       3442 STTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGG 3501
                :|||::|||||| |||||||||: |:|||||:|||||||||||||| :|| :|| |
Db        594 -CGTATATATCACAG-AAAAAACAGGGATAGCTCTAGGAGTCCTTACACATGTCTGAGAG 651

Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT-GGCAAAGGGTTG 3560
             |:|||||||| |||::|||::||||||:|||| |||:||||||| | ||| || || ||
Db        652 ACGAGCTTGC-ACCTGTGGCATACCTGAGTAAGAAAACTGATGTANTGGGCNAAAGGNTG 710

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
             :|:|||||| |||:||||| |||  ||| ||||||
Db        711 GCCTCATTGGTTATGGGTAATGGCAGCAATAGCAG 745


RESULT 126
ADB84407
ID   ADB84407 standard; DNA; 758 BP.
XX
AC   ADB84407;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #24.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Disclosure; Page 82; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
```

```
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 758 BP; 240 A; 186 C; 153 G; 168 T; 0 U; 11 Other;

  Query Match             7.4%;  Score 555.4;  DB 1;  Length 758;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  656;  Conservative   16;  Mismatches   83;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
Db        117 CGCNATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
Db        177 GATNTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
Db        237 TTCNCTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
Db        297 GCANTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
Db        353 GGCCCCGNTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
Db        413 CACTTACNCTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
Db        473 CTCAAAG---NTCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
Db        530 CCTATAACCTNACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
Db        590 GATGCTATATNGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
Db        650 CTTAGTTGCANACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
Db        710 AACCTGTCCCNGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758

RESULT 127
```

```
AAT37983
ID   AAT37983 standard; cDNA; 648 BP.
XX
AC   AAT37983;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   Rheumatoid arthritis associated retrovirus pol gene clone MSRV1nPR.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS   Retroviridae.
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 36-37; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is clone
CC   MSRV1nPR which corresponds to a fragment of the pol gene amplified in a
CC   nested PCR from retrovirus RNA isolated from synovial joint fluid cells
CC   of a patient with RPA. The sequences can be used for the diagnosis,
CC   prevention and treatment of RPA caused by infection or reactivation of
CC   the retroviruses. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 648 BP; 194 A; 156 C; 132 G; 163 T; 0 U; 3 Other;

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
```

```
                |||||||||:|||||||||||||||||:::|||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
                |||||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
                |||||||::|| :||||||||||:|||||:|||||:|||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
                |||||||||||:|||||||||:| ||||||:: ||   |||||||||| |  ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
                | |||  ||||||||:||||||||||:||||||||:|||||||||||||:||:|||::
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
                |||||| |||||| ||:||:||||||:||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
                |||||::||||::| ||||:||||||:::|||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 128
AAT37984
ID   AAT37984 standard; cDNA; 648 BP.
XX
AC   AAT37984;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   Rheumatoid arthritis associated retrovirus pol gene clone MSRV1.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS   Retroviridae.
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 37; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is clone MSRV1
CC   which corresponds to a fragment of the pol gene amplified in a nested PCR
CC   from retrovirus RNA isolated from sclerotic plaque cells of a patient
CC   with RPA. The sequences can be used for the diagnosis, prevention and
CC   treatment of RPA caused by infection or reactivation of the retroviruses.
CC   (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 648 BP; 194 A; 156 C; 132 G; 163 T; 0 U; 3 Other;

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
                || ||||:||||||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
                ||||||||||||||||||:|||: |||||||||||||||:|   ||::||||||||||||
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
                ||||   ||||||||| |||| :|||||:| |||||||  :|||| ||||||||:|||||
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
                |||||:|:|:||||||:||||||||||||||||||||||||| |||||:||||:||||||||
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
                |||||||||:|||||||||||||||:::|||||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
                |||||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
                |||||||::|| :|||||||||||:|||||:|||||:|||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
                |||||||||||:|||||||||:| ||||||:: ||   |||||||||| |  ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
                | |||  ||||||||:||||||||||:||||||||:|||||||||||||:||:|||::
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
                |||||| |||||| ||:||:||||||:||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
                |||||::||||::| ||||:||||||:::|||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 129
AAT37980
ID   AAT37980 standard; cDNA; 645 BP.
XX
AC   AAT37980;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   MSRV-1B POL gene fragment M003-P004.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
```

```
OS   Retroviridae.
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 26-27; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is clone M003-
CC   P004 which lies between clones MSRV-1B (AAT37974-9) and clone PSJ17
CC   (AAT37981) and corresponds to a fragment of the pol gene from the
CC   viruses. The sequences can be used for the diagnosis, prevention and
CC   treatment of RPA caused by infection or reactivation of the retroviruses.
CC   (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 645 BP; 191 A; 167 C; 131 G; 153 T; 0 U; 3 Other;

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             ||||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||  |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             | | |||  ||||||||:|||||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 130
AAT96523
ID   AAT96523 standard; cDNA; 645 BP.
XX
AC   AAT96523;
XX
DT   06-MAR-1998  (first entry)
XX
DE   MSRV-1B clone M003-P004.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; primer; PCR; ss;
KW   amplification; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Example 4; Page 96; 188pp; French.
XX
CC   This is the sequence of a clone named M003-P004 which is derived from the
CC   multiple sclerosis related virus 1B (MSRV-1B), and lies between the 3'
CC   end of a clone designated "pol MSRV-1B" and the 5' end of the isolated
CC   clone PSJ17 (AAT96517). The sequence was PCR amplified by primers
CC   AAT96518-T96522. MSRV-1 sequences are associated with multiple sclerosis
CC   and some are also associated with rheumatoid arthritis. Nucleic acid
CC   sequences from the MSRV-1 virus can be used in the diagnosis of multiple
CC   sclerosis and rheumatoid arthritis
XX
SQ   Sequence 645 BP; 191 A; 167 C; 131 G; 153 T; 0 U; 3 Other;

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
```

```
                |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
                |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:|||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
                |||||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
                :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
                || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
                ||||||||| |||||:|||||||||:|:||    :::|||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
                | | |||  |||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
                |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
                |||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645
```

RESULT 131
AAV43122
ID AAV43122 standard; cDNA; 645 BP.
XX
AC AAV43122;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus-1 clone MOO3-POO4 DNA sequence.
XX
KW Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN WO9823755-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-IB001482.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX
DR WPI; 1998-322732/28.
XX
PT New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT treatment of, e.g. multiple sclerosis.
XX
PS Example 4; Fig 7; 286pp; English.
XX
CC The present sequence represents the multiple sclerosis (MS) associated
CC retrovirus-1 (MSRV-1) clone MOO3-POO4 DNA sequence. The invention
CC provides complete or partial genomic sequences of the MSRV-1 pol gene,
CC gag gene and env gene, and polypeptides encoded by these genes. The
CC invention also provides antibodies raised against the polypeptides. The
CC genomic sequences, polypeptides and antibodies are also claimed useful
CC for diagnosing infection by MS and rheumatoid arthritis-associated
CC viruses, and also for prevention and treatment of infection with these
CC viruses
XX
SQ Sequence 645 BP; 191 A; 167 C; 131 G; 153 T; 0 U; 3 Other;

Query Match 7.0%; Score 523.9; DB 1; Length 645;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 550; Conservative 61; Mismatches 34; Indels 7; Gaps 4;

```
Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
                |||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
                |||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
                |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db         121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
                |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:|||||:|||||
Db         181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
                |||||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||||
Db         241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
                :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db         301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
                || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db         361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
                |||||||||  |||||:|||||||||:|:||    :::|||||||||||||||||||  ||
Db         420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
                | | |||  |||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db         477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
                |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db         535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
                |||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db         594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645
```

RESULT 132
ADB84556
ID ADB84556 standard; cDNA; 645 BP.
XX
AC ADB84556;
XX
DT 04-DEC-2003 (first entry)
XX
DE MSRV-1 DNA clone M003-P004.
XX
KW MSRV; ss; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
OS Multiple sclerosis associated retrovirus.

XX
PN US2003039664-A1.
XX
PD 27-FEB-2003.
XX
PF 26-NOV-1997; 97US-00979847.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Example 4; Fig 7; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is an MSRV cDNA clone. Note: The SEQ ID numbers for the
CC sequences as displayed in the main body of the patent do not match the
CC SEQ ID numbers in the sequence listing. Consequently those sequences
CC mentioned in the claims may not be the sequences the authors intended to
CC claim.
XX
SQ Sequence 645 BP; 191 A; 167 C; 131 G; 153 T; 0 U; 3 Other;

Query Match 7.0%; Score 523.9; DB 1; Length 645;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 550; Conservative 61; Mismatches 34; Indels 7; Gaps 4;

```
Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||| :|||||:| ||||||| |:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645
```

RESULT 133
ADG14749
ID ADG14749 standard; cDNA; 645 BP.
XX
AC ADG14749;
XX
DT 26-FEB-2004 (first entry)
XX
DE MSRV-1 clone PSJ17 PCR product.
XX
KW ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003198647-A1.
XX
PD 23-OCT-2003.
XX
PF 03-APR-2002; 2002US-00114104.
XX
PR 26-NOV-1996; 96US-00756429.
PR 26-NOV-1997; 97US-00979847.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2004-032461/03.
XX
PT New isolated nucleic acid and their fragments having the pol gene of a
PT retrovirus, useful for diagnosing, preventing and/or treating multiple
PT sclerosis and/or rheumatoid arthritis.
XX
PS Example 3; SEQ ID NO 8; 193pp; English.
XX
CC The invention relates to an isolated nucleic acid which comprises the pol
CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC arthritis. The methods and compositions of the present invention are

```
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 645 BP; 191 A; 167 C; 131 G; 153 T; 0 U; 3 Other;

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :||||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:|||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 134
AAT02822
ID   AAT02822 standard; cDNA; 645 BP.
XX
AC   AAT02822;
XX
DT   20-MAR-1996  (first entry)
XX
DE   Multiple sclerosis associated retrovirus clone MSRV-1B M003-P004.
XX
KW   Pol gene; multiple sclerosis; infectious agent; PCR; amplification; ss;
KW   retrovirus; ERV-9; HSERV-9; endogenous; MSRV-2; probe; primer; antibody.
XX
OS   Synthetic.
XX
PN   WO9521256-A1.
XX
PD   10-AUG-1995.
XX
PF   06-FEB-1995;   95WO-FR000142.
XX
PR   04-FEB-1994;   94FR-00001529.
PR   04-FEB-1994;   94FR-00001530.
PR   04-FEB-1994;   94FR-00001531.
PR   04-FEB-1994;   94FR-00001532.
PR   24-NOV-1994;   94FR-00014322.
PR   23-DEC-1994;   94FR-00015810.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Mallet F,  Mandrand B,  Bedin F,  Beseme F;
XX
DR   WPI; 1995-283776/37.
XX
PT   Two pathogenic or infectious agents associated with multiple sclerosis -
PT   one being a retrovirus, useful for early diagnosis, prevention and
PT   treatment.
XX
PS   Claim 3; Page 59; 104pp; French.
XX
CC   The partial nucleotide sequence of the pol gene from the multiple
CC   sclerosis associated infectious agent clone designated MSRV-1B M003-P004.
CC   The MSRV-1B (AAT02815-23) and MSRV-2 clones (AAT02824-6) are derived from
CC   the same strain of infectious agent, either POL2 or MS7PG or their
CC   variants, and can be produced by the same cell line, LM7PC or PLI-2 or
CC   other cell culture infected with the multiple sclerosis associated
CC   agents. The clones were isolated by extracting nucleic acids from cells
CC   of multiple sclerosis patients, converting the nucleic acid to cDNA and
CC   PCR amplifying a region of the pol gene conserved in retroviruses. The
CC   clones isolated fell into 2 distinct types: one was similar to the ERV-9
CC   and HSERV-9 endogenous retroviruses and designated MSRV-1B; the others
CC   were highly homologous to the known agent MSRV-2 and designated MSRV-2B.
CC   The nucleic acid, probes or primers derived from these sequences can be
CC   used for detecting agents associated with multiple sclerosis. The probes
CC   may be used to discriminate between these agents, opt. in conjunction
CC   with amplification primers and antibodies specific for the infectious
CC   agents
XX
SQ   Sequence 645 BP; 190 A; 167 C; 132 G; 153 T; 0 U; 3 Other;

  Query Match             7.0%;  Score 521.1;  DB 1;  Length 645;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  549;  Conservative   60;  Mismatches   36;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             | ||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TGAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :||||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||   :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTCCASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419
```

```
Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:||||||||||:|:||    :::||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             | | |||  ||||||||:||||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645
```

RESULT 135
ACH89452/c
ID ACH89452 standard; DNA; 591 BP.
XX
AC ACH89452;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #22647.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 22647; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 591 BP; 150 A; 134 C; 121 G; 186 T; 0 U; 0 Other;

Query Match 6.9%; Score 520.3; DB 1; Length 591;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 523; Conservative 40; Mismatches 28; Indels 1; Gaps 1;

```
Qy       3440 TTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAG 3499
             ||:||:|||::|||||||||||||||:||:| ||||:|||||||||||||||:|||:||
Db        591 TTCTTTATATGTCACAGAAAAAAACAGGAATAACTCTAGGAGTCCTTACACAGGTCCAAG 532

Qy       3500 GGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             | |:|||||||||||::||||: ||||||:||||||||:|||||||||||||||||||||
Db        531 GAACGAGCTTGCAACCCGTGGCAGACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 472

Qy       3560 GRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAA 3619
             |:|:|||||||||:||||| |||||||||||||| |::|||||||||||||||||||||
Db        471 GGCCTCATTGTTTATGGGTAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTTAAAA 412

Qy       3620 TAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTG 3679
             ||||||||| :|||||||||||||||||||||||||||||:|||||::||||||||||||
Db        411 TAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTG 352

Qy       3680 CTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTG 3739
             |||||||||||| |||||||||||||||| |: ||||||||||: |||| ||||||||||
Db        351 CTAAAGGAGACCTGTGGCTGTCAGACAGCCATTTACTTAAATACCAGGGTCTATTACTTG 292

Qy       3740 AARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAG 3799
             ||:|||||||||||:|||||||||||| |||||||||||||||||:| ||||||||||||
Db        291 AAGGGCCAGTGCTGCAACTGTGCACTTGGGCAACTCTTAACCCAGCCACATTTCTTCCAG 232

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
             ||| |||| ||||||||:||: |||||||||||| :| ||| |||||||||||||| |||
Db        231 ACATTGAAAAAAAGATAGAACGTAACTGTCAACAGGTGATTGCTCAAACCTATGCCGCTC 172

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
             |||| ||||:||||:||||:||||||||||:|||| |||||||||||| |||||||
Db        171 GAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGACC-TCAACTTGTATATTGATGGA 113

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
             |||||||||||||||| ||||||||||: |||||||||||||||||||||||||||||
Db        112 AGTTCCTTTGTAGAAAAAAGACTTCGAAAAGCAGGGTATGCAGTGGTCAGTGATAATGGA 53

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             |||:|||||| ||||||||||||||||||||||| |||:||||:||| |||
Db         52 ATACTTGAAAATAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAAAAC 1
```

RESULT 136
AAV43215
ID AAV43215 standard; cDNA; 635 BP.
XX
AC AAV43215;
XX
DT 29-DEC-1998 (first entry)

```
XX
DE   Multiple sclerosis associated retrovirus fragment 4.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .234
FT                   /*tag=  a
FT                   /product= "Encodes protein AAW71067"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71067.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 183; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) genomic fragment used in the method of the invention.
CC   The invention provides complete or partial genomic sequences of the MSRV-
CC   1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC   genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
             |||||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
             |||||||||||||:||||||| | || |||||||||||||  |||||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
             ||||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
             : || ||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||  |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
             ||||||||| |||||| |||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
             |||||||||||||||||||||| | ||||||||||||||||||||||||:|||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
             |||||||||| |||||||||| |||||||||||||||||| | ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
             |||||||||||||||||||||||||||||||| |||||||||||  |||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
             |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 137
AAX29702
ID   AAX29702 standard; DNA; 635 BP.
XX
AC   AAX29702;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Clone CL6-3' from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis; ss.
XX
OS   Viruses.
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   P-PSDB; AAW99552.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 1; Page 36-37; 83pp; French.
XX
CC   This sequence represents clone CL6-3' from a novel multiple sclerosis
CC   related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC   prophylactic or therapeutic compositions to inhibit expression of a
CC   multiple sclerosis related virus and/or virus associated with rheumatoid
CC   polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC   17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
             |||||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60
```

```
Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:|||||| | || |||||||||||||  |||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              |||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              |||||||||||||||||||||  |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||| |||||| ||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | ||||||||||||||||||||||||:||||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617
```

RESULT 138
ADG14843
ID  ADG14843 standard; cDNA; 635 BP.
XX
AC  ADG14843;
XX
DT  26-FEB-2004  (first entry)
XX
DE  MSRV associated cDNA #7.
XX
KW  ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS  Multiple sclerosis associated retrovirus.
XX
PN  US2003198647-A1.
XX
PD  23-OCT-2003.
XX
PF  03-APR-2002; 2002US-00114104.
XX
PR  26-NOV-1996;  96US-00756429.
PR  26-NOV-1997;  97US-00979847.
XX
PA  (INMR ) BIO MERIEUX.
XX
PI  Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI  Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR  WPI; 2004-032461/03.
XX
PT  New isolated nucleic acid and their fragments having the pol gene of a
PT  retrovirus, useful for diagnosing, preventing and/or treating multiple
PT  sclerosis and/or rheumatoid arthritis.
XX
PS  Disclosure; SEQ ID NO 102; 193pp; English.
XX
CC  The invention relates to an isolated nucleic acid which comprises the pol
CC  gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC  arthritis. The methods and compositions of the present invention are
CC  useful for diagnosing, preventing and/or treating multiple sclerosis
CC  and/or rheumatoid arthritis. The present sequence is used in the
CC  exemplification of the invention.
XX
SQ  Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;

```
  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
              |||||||||||:|||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:|||||| | || |||||||||||||  |||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              |||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              |||||||||||||||||||||  |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||| |||||| ||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | ||||||||||||||||||||||||:||||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617
```

RESULT 139
AAX25685
ID  AAX25685 standard; DNA; 591 BP.
XX
AC  AAX25685;
XX
DT  21-MAY-1999  (first entry)
XX
DE  Human endogenous retrovirus W probe Penv-C15.
XX

```
KW   Clone; human endogenous retrovirus; genome; autoimmune disease; probe;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Synthetic.
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Example 6; Page 92; 106pp; French.
XX
CC   This sequence represents a probe used to detect tissue distribution and
CC   copy number of the human endogenous retrovirus (HERV) W genome (AAX25665)
CC   in the human genome. Nucleic acids, their fragments or peptides encoded
CC   by them derived from the HERV-W genome are markers of autoimmune disease
CC   (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus
CC   erythematosus, insulin-dependent diabetes and related pathologies) and of
CC   abnormal or unsuccessful pregnancy and can be used as chromosomal markers
CC   for susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 591 BP; 177 A; 141 C; 113 G; 160 T; 0 U; 0 Other;

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              |||||||||||||||||||||||||||  ||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||:|||||||||||||||||:||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
               || |||||||||||||||:|||||||||||||| |||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| ||||||||||||||||||||||||:|||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||| |||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 140
AAA59204
ID   AAA59204 standard; DNA; 591 BP.
XX
AC   AAA59204;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Probe Penv-C15 used to isolate a fragment of HERV-W.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 42; 53pp; French.
XX
CC   The present sequence represents a probe which is used to detect and
CC   isolate an endogenous retroviral nucleic acid fragment, which is
CC   associated with an autoimmune disease, and is integrated into the human
CC   genome. The fragment is originally derived from a novel retrovirus, human
CC   endogenous retrovirus W (HERV-W). The fragment comprises (or consists of)
CC   at least part of the gag gene of the HERV-W retrovirus, which is
CC   associated with either autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 591 BP; 177 A; 141 C; 113 G; 160 T; 0 U; 0 Other;

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              |||||||||||||||||||||||||||  ||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120
```

```
Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||||:||||||||||||||||||:|||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||||| |||||| ||||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              || |||||||||||||||:||||||||||||| ||||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| |||||||||||||||||||||||:|||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576
```

RESULT 141
AAV09124
ID AAV09124 standard; cDNA; 542 BP.
XX
AC AAV09124;
XX
DT 28-MAY-1998 (first entry)
XX
DE Nucleotide sequence encoding the AJ172_2 protein.
XX
KW AJ172_2 clone; cDNA library; human adult testes; cell proliferation;
KW cell differentiation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 441. .542
FT /*tag= a
FT /product= "AJ172_2 protein"
XX
PN US5708157-A.
XX
PD 13-JAN-1998.
XX
PF 26-JUL-1996; 96US-00686878.
XX
PR 26-JUL-1996; 96US-00686878.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Mccoy JM, Treacy M, Racie LA, Merberg D, Spaulding V, Jacobs K;
PI Lavallie ER, Evans C;
XX
DR WPI; 1998-109352/10.
DR P-PSDB; AAW42030.
XX
PT Human testes derived cDNA and protein(s) - may be useful for laboratory
PT use, or for stimulating cell proliferation and differentiation.
XX
PS Disclosure; Col 81-82; 48pp; English.
XX
CC This nucleotide sequence encodes the AJ172-2 protein, which is 67 amino
CC acids in length, and is isolated from a human adult testes cDNA library.
CC When the AJ172_2 sequence was searched in the Genbank database, the clone
CC demonstrated at least some homology with the following sequences: Friend
CC murine leukemiz virus; Moloney murine leukemia virus genome; pol protein,
CC Gibbon leukemia virus; and an EST identified as "yh46a09.s1 Homo sapiens
CC cDNA clone 132760 3'"
XX
SQ Sequence 542 BP; 150 A; 148 C; 93 G; 141 T; 0 U; 10 Other;

Query Match 6.8%; Score 510.3; DB 1; Length 542;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 516; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

```
Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
              ||||| || |||||||||||||||||| ||||| |||| |||||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
              ||||::|||||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
              | ||||| ||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
              |||||||||||||||||||||||||||||||||||||||| ||||||||:|||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCCTCCATGCCGC 5658
              ||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
              |||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
              |||
Db        540 CAT 542
```

RESULT 142
AAV15134
ID AAV15134 standard; cDNA; 542 BP.
XX
AC AAV15134;
XX
DT 29-JUN-1998 (first entry)
XX
DE Human adult testes cDNA clone AJ172_2 internal nucleotide sequence.
XX
KW Human; adult; testes; AJ172_2; secreted protein; cytokine; nutrition;
KW cell proliferation; differentiation; tumour; immune; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

```
FT   CDS             342. .542
FT                   /*tag=  a
FT                   /note= "no stop codon given"
XX
PN   WO9804696-A1.
XX
PD   05-FEB-1998.
XX
PF   23-JUL-1997;   97WO-US012910.
XX
PR   26-JUL-1996;   96US-00686878.
PR   27-SEP-1996;   96US-00721489.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Racie LA,  Merberg D,  Treacy M;
PI   Spaulding V;
XX
DR   WPI; 1998-130687/12.
DR   P-PSDB; AAW44290.
XX
PT   Polynucleotides and secreted proteins having cytokine and cell
PT   proliferation/differentiation activity - isolated from a human adult
PT   testes cDNA library.
XX
PS   Claim 1; Page 33; 48pp; English.
XX
CC   The present sequence represents the internal nucleotide sequence of clone
CC   AJ172_2 from a human adult testes cDNA library. The nucleotide sequences
CC   and protein from clone AJ172_2, have uses in nutrition, cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis regulating activity, chemotactic/chemokinetic
CC   activity, haemostatic or thrombotic activity, receptor/ligand activity,
CC   anti-inflammatory activity, tumour inhibition activity or other
CC   activities. The products can also be used for production of antibodies
CC   and for detection and diagnosis
XX
SQ   Sequence 542 BP; 150 A; 148 C; 93 G; 141 T; 0 U; 10 Other;

 Query Match             6.8%;  Score 510.3;  DB 1;  Length 542;
 Best Local Similarity   95.0%;  Pred. No. 0;
 Matches  516;  Conservative   12;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
             ||||| || |||||||||||||||||||| ||||| |||| |||||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
             ||||::|||||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
             | ||||| ||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
             |||||||||||||||||||||||||||||||||||| |||||||:|||||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
             ||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
             |||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
             |||
Db        540 CAT 542


RESULT 143
ADB84397
ID   ADB84397 standard; DNA; 635 BP.
XX
AC   ADB84397;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #17.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Claim 31; Page 76-77; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
```

```
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is a claimed MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 635 BP; 188 A; 167 C; 133 G; 138 T; 0 U; 9 Other;

  Query Match             6.8%;  Score 510;  DB 1;  Length 635;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  544;  Conservative   11;  Mismatches   50;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
             ||||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  ||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCN 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
             |||||||||||||:|||||| | || |||||||||||||||  |||||||||||||| ||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAN 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
             |||||||| |||||||||| |||||| |||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGN 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
             : || ||||||||||||||||||||||| |||||||||||||||| ||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAN 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||||  |: ||||||||:|||||||||:|||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACN 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||| |||||| |||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTN 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
             |||||||||||||||||||| | ||||||||||||||||||||||:|||||||:|
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTN 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
             |||||||||| |||||||||| |||||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGN 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
             |||||||||||||||||||||||||||||||| ||||||||||||  ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCN 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 144
AAX25683
ID   AAX25683 standard; DNA; 678 BP.
XX
AC   AAX25683;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human endogenous retrovirus W probe Ppol-MSRV.
XX
KW   Clone; human endogenous retrovirus; genome; autoimmune disease; probe;
KW   multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW   disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS   Synthetic.
OS   Human endogenous retrovirus.
XX
PN   WO9902696-A1.
XX
PD   21-JAN-1999.
XX
PF   06-JUL-1998;   98WO-FR001442.
XX
PR   07-JUL-1997;   97FR-00008815.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Beseme F,  Blond J,  Bouton O,  Mandrand B,  Mallet F;
XX
DR   WPI; 1999-120897/10.
XX
PT   New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT   exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT   disease, and abnormal or failed pregnancy.
XX
PS   Example 6; Page 90-91; 106pp; French.
XX
CC   This sequence represents a probe used to detect tissue distribution and
CC   copy number of the human endogenous retrovirus (HERV) W genome (AAX25665)
CC   in the human genome. Nucleic acids, their fragments or peptides encoded
CC   by them derived from the HERV-W genome are markers of autoimmune disease
CC   (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus
CC   erythematosus, insulin-dependent diabetes and related pathologies) and of
CC   abnormal or unsuccessful pregnancy and can be used as chromosomal markers
CC   for susceptibility to these conditions, or proximity markers of genes
CC   associated with this susceptibility
XX
SQ   Sequence 678 BP; 202 A; 171 C; 142 G; 162 T; 0 U; 1 Other;

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| |||||||||||||||:||:||||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:|||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||              ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:|||||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||
```

```
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||||:|||:||||||: |||||||||||::||||::|||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
             |:|||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 145
AAA59203
ID   AAA59203 standard; DNA; 678 BP.
XX
AC   AAA59203;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Probe Ppol-MSRV used to isolate a fragment of HERV-W.
XX
KW   Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW   gag gene; pregnancy; multiple sclerosis; T cell proliferation; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043521-A2.
XX
PD   27-JUL-2000.
XX
PF   21-JAN-2000; 2000WO-FR000144.
XX
PR   21-JAN-1999;   99FR-00000888.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Paranhos-Baccala G,  Mallet F,  Voisset C;
XX
DR   WPI; 2000-499229/44.
XX
PT   New nucleic acid from human endogenous retrovirus, useful e.g. for
PT   diagnosis of autoimmune disease and complications of pregnancy, contains
PT   at least part of the gag gene.
XX
PS   Disclosure; Page 41-42; 53pp; French.
XX
CC   The present sequence represents a probe which is used to detect and
CC   isolate an endogenous retroviral nucleic acid fragment, which is
CC   associated with an autoimmune disease, and is integrated into the human
CC   genome. The fragment is originally derived from a novel retrovirus, human
CC   endogenous retrovirus W (HERV-W). The fragment comprises (or consists of)
CC   at least part of the gag gene of the HERV-W retrovirus, which is
CC   associated with either autoimmune disease, failure of pregnancy or
CC   disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC   from it, are useful for diagnosis of autoimmune disease (specifically
CC   multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC   fragments may also be used for in situ labelling of isolated chromosomes,
CC   while the transcription product can be used to study or monitor T cell
CC   proliferation in vitro
XX
SQ   Sequence 678 BP; 202 A; 171 C; 142 G; 162 T; 0 U; 1 Other;

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| |||||||||||||||:||:|||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG-----------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||           ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:|||||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||| ||||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||   ||  |   |||  ||||||||:||||||||||
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||||:|||:|||||: |||||||||||::||||::|||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
             |:|||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 146
AAS88391
ID   AAS88391 standard; cDNA; 583 BP.
XX
AC   AAS88391;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #24195.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
```

```
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG24204.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 24195; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 583 BP; 174 A; 156 C; 133 G; 120 T; 0 U; 0 Other;

  Query Match             6.5%;  Score 489.8;  DB 1;  Length 583;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  515;  Conservative   11;  Mismatches   28;  Indels   10;  Gaps    1;

Qy       7019 TCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 7078
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 60

Qy       7079 GACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 7138
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 120

Qy       7139 CTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGA 7198
              |||||||||||||||||||||||| |||| | |||||| |||||||||||||||| ||||
Db        121 GTGAGGAAATCTCAGCTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGA 180

Qy       7199 GCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGA 7258
              |||||:||| | |||||||||||| |||||||| |||||||||||||||||| |||| |||||
Db        181 GCGGTCGTCAGTCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGA 240

Qy       7259 GAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGG 7318
              |||||||||||||||||||||||||||| ||:|||||||||:|||||||||:||||||||
Db        241 GAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGG 300

Qy       7319 TGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAG 7378
              ||||| | |||||||||||||||||||||||||||||||:|||||||| ||||||||||||
Db        301 TGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAG 360

Qy       7379 AGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATT 7438
              |||||||||||||||||||| ||||| |||||||||||||||||||||| ||||:|||
Db        361 AGCTCACTAAAATGCTAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATC 420

Qy       7439 GCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGC 7498
              ||:|||||||||||| |||||||||||||||||||||||||| |  |||||||| ||
Db        421 GCCTGAGAGCACAGCAAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGC 480

Qy       7499 AACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTC 7558
              |||||| |||| |||||||||||||||||||||||||||||| ||||          ||||
```

```
Db        481 AACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTC 530

Qy       7559 ACTCTATTAAATCTTGCARCTGCR 7582
              ||||||||||||||||||:||||:
Db        531 ACTCTATTAAATCTTGCAACTGCA 554


RESULT 147
ABA61335/c
ID   ABA61335 standard; DNA; 549 BP.
XX
AC   ABA61335;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #9640.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 9640; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              |||||||||::|||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
              |||||||:|||||||||  ||||||||||::|||||||||||||  ||||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
              |||:||||||||||  |||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370
```

```
Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             ||||||||||||||||||||||: ||||||||||:||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1
```

RESULT 148
AAI41239/c
ID AAI41239 standard; DNA; 549 BP.
XX
AC AAI41239;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #9925 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 9925; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

Query Match 6.4%; Score 481.1; DB 1; Length 549;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 488; Conservative 33; Mismatches 28; Indels 1; Gaps 1;

```
Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::||||||||||||  |||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  ||||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: ||||||||||:||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1
```

RESULT 149
ABA29141/c
ID ABA29141 standard; DNA; 549 BP.
XX
AC ABA29141;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #7607 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.

```
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 7607; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: |||||||||||:||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| ||||||||||||||||||||:| ||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||| |: | |||||||||||| ||||| ||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| |||||:||||:||||||||||||:|||:|
Db         70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1


RESULT 150
AAK35525/c
ID   AAK35525 standard; DNA; 549 BP.
XX
AC   AAK35525;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 10082.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 10082; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: |||||||||||:||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
```

```
            ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
            |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
            ||||:||||:||||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
            |||||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
            |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1
```

RESULT 151
AAK09633/c
ID AAK09633 standard; DNA; 549 BP.
XX
AC AAK09633;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 9624.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 9624; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX

SQ Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

Query Match 6.4%; Score 481.1; DB 1; Length 549;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 488; Conservative 33; Mismatches 28; Indels 1; Gaps 1;

```
Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
            |||||||||::||||::|.||||:||||||||:||||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
            |||||||:|||||||||  ||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
            |||:|||||||||||  ||||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
            |||||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
            ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
            |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
            ||||:||||:||||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
            |||||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
            |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1
```

RESULT 152
ABS35255/c
ID ABS35255 standard; DNA; 549 BP.
XX
AC ABS35255;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 10245.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.

```
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 10245; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:||||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  ||||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  ||||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: ||||||||||||:|||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: ||||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:||||:||||||||||: ||| ||||||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| |||||||:||||:|||||||||||||||:|||:|
Db         70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
```

```
                 |||||||||
Db          10 AGGCACTAGA 1
```

```
RESULT 153
ABS09860/c
ID   ABS09860 standard; DNA; 549 BP.
XX
AC   ABS09860;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 9851.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 9851; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
```

```
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 549 BP; 139 A; 124 C; 120 G; 166 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||| ||||||||||::|||||||||||| ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:|||||||||  |||||||||||||||||:|||||::|||||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: |||||||||||:||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:|||| ||| |||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1


RESULT 154
ACH75729/c
ID   ACH75729 standard; DNA; 559 BP.
XX
AC   ACH75729;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #8924.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 8924; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 559 BP; 118 A; 125 C; 121 G; 195 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 477.3;  DB 1;  Length 559;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  489;  Conservative   33;  Mismatches   37;  Indels    1;  Gaps    1;

Qy       3840 TTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCT 3899
             | |||||||||||||| |||||||  |||||:||||:|||||:||||||||||:||||
```

```
Db        559 TGCTCAAACCTATGCCGCTCGAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGACC- 501

Qy       3900 TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATG 3959
              |||||||||||| |||||||||||||||||||||||| ||||||||||||: |||||||
Db        500 TCAACTTGTATATTGATGGAAGTTCCTTTGTAGAAAAAAGACTTCGAAAAGCAGGGTATG 441

Qy       3960 CAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTY 4019
              |||||||||||||||||||||||:|||||| |||||||||||||||||||||||| |||:
Db        440 CAGTGGTCAGTGATAATGGAATACTTGAAAATAATCCCCTCACTCCAGGAACTAGCGCTC 381

Qy       4020 AGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA 4079
              ||||:||| ||||||||||:| |:| |||||||| |||||| | |||:||||||||:||
Db        380 AGCTGGCAAAACTAATAGCCCTTACTCGGGCACTAAAATTAGAAAAAGAAAAAAGGGTAA 321

Qy       4080 ATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSA 4139
              | |||||||||||||||:|:|||||||:|||||||| |||||||||||::|||||||:|
Db        320 ACATATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGA 261

Qy       4140 RAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGA 4199
              :|||||||||||| |||||||||:|||:||||| ||||||:|||||||||||||| ||||
Db        260 GAGAAAGGGAATACCTAACTTCTGAGGGAACATCTATCAAACATCAGGAAGCCATCAGGA 201

Qy       4200 RATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATC 4259
              :|||||||: |||:|||||||||||||:||||||||:|||||||||||||:|||||||||
Db        200 GATTATTATCGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATC 141

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
              | |||||||||:|||||||||||:|||:|||: ||||||| ||||||||||:|||||||
Db        140 AGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCGAATATTGAAGCCAAAAGAG 81

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
              |  |||||||||||||||||||| |||||||||| ||   ||| |||||| ||||||||
Db          80 CCACAAGGCAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATC 21

Qy       4380 CCTTCCGGGAAACCAAGCCC 4399
              || || ||||||||||||||
Db          20 CCCTCTGGGAAACCAAGCCC 1
```

RESULT 155
ACH67368/c
ID ACH67368 standard; DNA; 568 BP.
XX
AC ACH67368;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #563.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 563; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 568 BP; 155 A; 109 C; 158 G; 146 T; 0 U; 0 Other;

Query Match 6.3%; Score 473.9; DB 1; Length 568;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 499; Conservative 30; Mismatches 36; Indels 3; Gaps 2;

```
Qy       2644 GCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGG 2703
              | |||||||||||  || ||||||| || |||||||||||||||||||||| |||| |
Db        565 GTACAGAAACCCAACGGGCAGTGGA-GGCTAGTGCAAGATCTCAGGATTATAAATGAGGT 507

Qy       2704 CCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAG 2763
                |||||| | ||||| |||||||||||| |||||||||| ||::|||||||||||||||
Db        506 A--TTGTCCCTCTATACTCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAG 449

Qy       2764 AGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTAC 2823
              |||||||||||||||||||:||||||||||:|||||||||| |||||||| ||||||
Db        448 AGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTTCTGCATCTCTGTAC 389

Qy       2824 ATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCA 2883
              | ||| |||||||||||||||||||||||||||| |||| ||||||:| ||||||||||
Db        388 AACCTAACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCGAACCCAACGTCTCAACTCA 329

Qy       2884 CCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAG 2943
              ||||||||:|||||||||||||||| |||||||: |||||||||||||||||||| |||
Db        328 CCTGGACTGTTTTACCCCAAGGGTTCCGGGATAGCTCCCATCTATTTGGCCAGGCGTTAA 269

Qy       2944 CCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATT 3003
              |||||||||||||:||:|::|||||||| ||||||||||||||||:|||:||||||||||
Db        268 CCCAAGACTTGAGCCAGTTCTCATACCTTGACACTCTTGTCCTTCAGTATGTGGATGATT 209

Qy       3004 TACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATT 3063
              | |||||:||: ||::|||||||||||||||||||||||| ||||||||:||||| ||||
Db        208 TGCTTTTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCTACCCAAGCGCTCTTGAATT 149
```

```
Qy       3064 TCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGG 3123
             |||||||:|||||||||||||:||||||||||| |:|:|||||:|||||||||||| ||||
Db        148 TCCTCGCCACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAACAGG 89

Qy       3124 TTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCC 3183
             |||||| ||||||:||||:||||||||||||||||||:|||||||||| |||||:::|||| 
Db         88 TTAAATGCTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTAAGGAATGTATCC 29

Qy       3184 AGCCTATACTGGCTTATCCTCATCYCAA 3211
             |||| |||||||||||||||||:|||
Db         28 AGCCCATACTGGCTTATCCTCATCCCAA 1


RESULT 156
AAI44528
ID   AAI44528 standard; DNA; 568 BP.
XX
AC   AAI44528;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #13214 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 13214; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 568 BP; 148 A; 162 C; 104 G; 154 T; 0 U; 0 Other;

  Query Match             6.2%;  Score 460.7;  DB 1;  Length 568;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  494;  Conservative   28;  Mismatches   31;  Indels    3;  Gaps    3;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |||||||||| |||||||||| ||||||||||||||||||||||||||
Db         16 ATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTAT 74

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             || | |||||||||||||| | |||||||||||||||||| |||||||||| ||::|||||
Db         75 CAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCC 133

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||| |||||||||||||||||| | ||||||||||:|||||||||| |||||||
Db        134 AAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATGCCTTTTTCTGCA 193

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||||||||||||||| |||||| |||||||||||||||| |||| ||||||:| |
Db        194 TCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTTCGAACCCAACGT 253

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||| |||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        254 CTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 313

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             || |||||||||||||||||||:||:|::|||||||||||||||||||||||||:||| |
Db        314 AGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCAGTAAG 373

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             || |||||||||||| :||:|||::|||||||||||||||||||||||||||||||||:|
Db        374 TGAATGATTTACTTTCAGCTGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCAC 433

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:|||||||| ||:|||||||||||:|||||||||:| |:|||||:|||||||
Db        434 TCTTAAATTTCCTTGCCACCTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCT 492

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             || | |||| ||||||||||||:||||:||||||||||||||||:|||||||||||||||
Db        493 CAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 552

Qy       3175 AAYRYATCCAGCCTAT 3190
             ||:::|||||||| ||
Db        553 AACGTATCCAGCCCAT 568


RESULT 157
ABA62059/c
ID   ABA62059 standard; DNA; 548 BP.
XX
AC   ABA62059;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #10364.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 10364; 639pp + Sequence Listing; English.
XX
```

```
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| ||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||::||| ::|||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||:|||||||| |||||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 158
AAI41998/c
ID   AAI41998 standard; DNA; 548 BP.
XX
AC   AAI41998;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #10684 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 10684; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| ||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||::||| ::|||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||:|||||||| |||||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
```

```
Db         12 CTTAACCCAGCC 1


RESULT 159
AAK36276/c
ID   AAK36276 standard; DNA; 548 BP.
XX
AC   AAK36276;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 10833.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 10833; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| |||||||||||||||||||||||||||||||| ||:|||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || |||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||||| |||:|| :|||||:|||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:||||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||:||||||||| ||||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||: ||||||||||:|||| ||||||||||||||||
Db         72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db         12 CTTAACCCAGCC 1


RESULT 160
AAK10373/c
ID   AAK10373 standard; DNA; 548 BP.
XX
AC   AAK10373;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 10364.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 10364; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
```

```
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|| :|||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||: |||| ||||||||||||||||||||||||||||||||| ||:|||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
              ||||:|||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:|||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||||||||||||||||||||||||||||||:|||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||:|||||||||||||||||||: |||||||||||:|||| ||||||||||||||
Db         72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
              ||||||||||:|
Db         12 CTTAACCCAGCC 1


RESULT 161
ABS35955/c
ID   ABS35955 standard; DNA; 548 BP.
XX
AC   ABS35955;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 10945.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 10945; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|| :|||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||: |||| ||||||||||||||||||||||||||||||||| ||:|||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
              ||||:|||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:|||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||||||||||||||||||||||||||||||:|||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||:|||||||||||||||||||: |||||||||||:|||| ||||||||||||||
Db         72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
```

```
                ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 162
ABS10362/c
ID   ABS10362 standard; DNA; 548 BP.
XX
AC   ABS10362;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 10353.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 10353; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 548 BP; 137 A; 129 C; 117 G; 165 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||: |||| ||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
              ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:|||||||||||| |||:|| :|||||:|||||||||||::||| ::|||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:||||::|||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||:|||||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
              ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 163
ACH77079
ID   ACH77079 standard; DNA; 578 BP.
XX
AC   ACH77079;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #10274.
```

```
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 10274; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 578 BP; 190 A; 120 C; 128 G; 140 T; 0 U; 0 Other;

  Query Match             6.0%;  Score 448.6;  DB 1;  Length 578;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  489;  Conservative   47;  Mismatches   42;  Indels    8;  Gaps    4;

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||: || :||| || ||||:|| ||:||||||:    |||| ||||||||||

Db             1 TTTCCCAGGTACGGTGAAACAGTCAGGCCAGTATATACAC----TAAGTAAACTCAGAAA 56

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||
Db            57 GCCAATACCCATTTAGTAAGATGGACACCTGAAGCAAAAGAGGCTTTCCAGGCCCTAAAG 116

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||| |||||||:|||||||||:|||||||| ||| |:||:||
Db           117 AAGGCCCT--AACCCAAGCCCCAATGTTAAGTTTGCCAACAGGGCAAGGCTTCTCTTTAT 174

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |::||||||||| ||||||:||:||||||:||| |||||||||||:| |:|||||: |||
Db           175 ATGTCACAGAAA-AAACAGGAATAGCTCTAGGAATCCTTACACAGGTTCAAGGGACCAGC 233

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
              ||||||||::||||::|| |||:|||| |||:|||||||||||||||||||||:|:||||
Db           234 TTGCAACCTGTGGCATACTTGAGTAAGAAAATTGATGTAGTGGCAAAGGGTTGGCCTCAT 293

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
              ||||||: |||||||  |||||||||||::|||||||||||||||||||||||||||||
Db           294 TGTTTACAGGTAGTGACGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAG 353

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:||||||||||||||||||||||||||||:|| ||::| ||||||||||||||||||
Db           354 GGAAGAGATCTTACTGTGTGGACATCTCATGATGTAAATGGTATACTCACTGCTAAAGGA 413

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||:|||||: ||| |||||| :||||||||||||| |||||:| | |
Db           414 AACTTGTGGCTGTCAGACAACCATTTGCTTAAAAATCAGGCTCTATTAATTGAAGGACTA 473

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||| :|||| |||||||||||||||||||||||:| |||||||||||||||||||
Db           474 GTGCTGAGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 533

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATG 3853
              ||||||| |:|.|: |||||||||||  :| ||| ||||||||||
Db           534 GAAAAGAGAGAACGTAACTGTCAAC-GGTGATTGCTCAAACCTATG 578


RESULT 164
AAI17000/c
ID   AAI17000 standard; DNA; 559 BP.
XX
AC   AAI17000;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #6933 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
```

```
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 6933; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| ||||||||||||||||||||||||||:|||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||            |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||| ||| |||||||| ||||| |||||||| || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |        ||    |||||||||||||||  ||| |||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||||  ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 165
ABA61156/c
ID   ABA61156 standard; DNA; 559 BP.
XX
AC   ABA61156;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #9461.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 9461; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| ||||||||||||||||||||||||||:|||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||            |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||| ||| |||||||| ||||| |||||||| || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140
```

```
Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |        ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                |||||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 166
AAI41056/c
ID   AAI41056 standard; DNA; 559 BP.
XX
AC   AAI41056;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #9742 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 9742; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
                |||||  |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
                |:||||||| |||||||||||||||||||||||||:||||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
                |||||||||||||  |||||||||||||||||||||||||||:|| |||          |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                |||||||||| ||| ||||||||  ||||| ||||||||  || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :|||||| | |||||||| |||||||  ||||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |        ||    |||||||||||||||  ||| |||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                |||||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 167
ABA29035/c
ID   ABA29035 standard; DNA; 559 BP.
XX
AC   ABA29035;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #7501 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 7501; 530pp; English.
```

```
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| ||||||||||||||||||||||||||||:||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             |||||||||||||  |||||||||||||||||||||||||:|| |||             |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  ||| ||||||||  |||||  ||||||||  ||  ||||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |        ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| ||||||||||| ||||||||||||
Db          86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||||:||||
Db          26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 168
AAK35343/c
ID   AAK35343 standard; DNA; 559 BP.
XX
AC   AAK35343;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 9900.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 9900; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| ||||||||||||||||||||||||||||:||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             |||||||||||||  |||||||||||||||||||||||||:|| |||             |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  ||| ||||||||  |||||  ||||||||  ||  ||||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |        ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
```

```
                ||||||||||||:|||||||||:|||||||||:| ||||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                ||||||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 169
AAK09451/c
ID   AAK09451 standard; DNA; 559 BP.
XX
AC   AAK09451;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 9442.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 9442; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
                ||||| |||||||||| |||||||||:|||||||||:|||| | |||||||| ||:|||||
Db         559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
                |:||||||| |||||||||||||||||||||||||||:||||||||||||||||||||
Db         499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
                |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:

Db         439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------C 7036
                |||||||||||||  |||||||||||||||||||||||||||:|| |||          |
Db         379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                |||||||||||| ||| |||||||||  ||||| |||||||||  || ||||||||| |||||
Db         319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||| |||
Db         259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||||
Db         199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |       ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db         139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                ||||||||||||:|||||||||:|||||||||:| ||||||||||| ||||||||||||
Db          86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                ||||||||||||||||||||||:||||
Db          26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 170
ABS35081/c
ID   ABS35081 standard; DNA; 559 BP.
XX
AC   ABS35081;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 10071.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 10071; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
```

```
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| |||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||             |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  ||| ||||||||  ||||| |||||||||  || |||||||||  |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||||  |||||||  |||||||||||||||||||||||||||  |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |       ||    |||||||||||||||  ||| |||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 171
ABS09734/c
ID   ABS09734 standard; DNA; 559 BP.
XX
AC   ABS09734;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 9725.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 9725; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
```

```
XX
SQ   Sequence 559 BP; 135 A; 120 C; 153 G; 151 T; 0 U; 0 Other;

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| |||||||||||||||||||||||||:||||||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||            |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  ||| ||||||||  ||||| ||||||||  || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||||  |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |     ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 172
AAC17319
ID   AAC17319 standard; cDNA; 443 BP.
XX
AC   AAC17319;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 21394.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 21394; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 443 BP; 83 A; 137 C; 94 G; 128 T; 0 U; 1 Other;

  Query Match             5.8%;  Score 436.8;  DB 1;  Length 443;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  435;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         12 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 62

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db         63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTKGCAACTGCACTCTTCTGGTC 122

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 182

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 242

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||| ||||
Db        243 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 302

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        303 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 362

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        363 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 422

Qy        432 TCCATTCCTTGGAATCCGTGA 452
              |||||||||||||||||||||
Db        423 TCCATTCCTTGGAATCCGTGA 443


RESULT 173
AAI49732
ID   AAI49732 standard; DNA; 491 BP.
XX
AC   AAI49732;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #18418 used to measure gene expression in human placenta sample.
```

```
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 18418; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 491 BP; 149 A; 104 C; 113 G; 125 T; 0 U; 0 Other;

  Query Match             5.8%;  Score 433.2;  DB 1;  Length 491;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  423;  Conservative   38;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
Db          1 AGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAGAAAAAAACAGGAATAACTCT 60

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
Db         61 AGGAGTCCTTACACAGGTCCAAGGAACGAGCTTGCAACCCGTGGCAGACCTGAGTAAGGA 120

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
Db        121 AATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTGGCAGTAGCAAT 180

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
Db        181 CTTAGTATCTGAAGCAGTTAAAATAATACAGGAAAGAGATCTTACTGTGTGGACATCTCA 240

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
Db        241 TGATGTGAACAGCATACTCACTGCTAAAGGAGACCTGTGGCTGTCAGACAGCCATTTACT 300

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
Db        301 TAAATACCAGGGTCTATTACTTGAAGGGCCAGTGCTGCAACTGTGCACTTGGGCAACTCT 360

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
Db        361 TAACCCAGCCACATTTCTTCCAGACATTGAAAAAAAGATAGAACGTAACTGTCAACAGGT 420

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
Db        421 GATTGCTCAAACCTATGCCGCTCGAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGA 480

Qy       3897 CCT 3899
Db        481 CCT 483


RESULT 174
ACH87164
ID   ACH87164 standard; DNA; 566 BP.
XX
AC   ACH87164;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #20359.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 20359; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
```

```
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 566 BP; 162 A; 152 C; 112 G; 140 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 429.5;  DB 1;  Length 566;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  507;  Conservative   11;  Mismatches   46;  Indels   13;  Gaps    5;

Qy       2411 GAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGC 2470
             ||||||||||||| ||||| |||||||||||||||||||||||||||||| ||||||||||
Db          1 GAAGGAATTAATCTTGAAGACTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGC 59

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             ||||||||||||||||||||||||||||  | |||||||| ||||||||||||||||||
Db         60 CCGTCCTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCCACCAAAGGCAGTACCC 119

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGG 2590
             ||| |||||| |||| ||||||||| |||||||||   |||||||| ||||||||||||
Db        120 CCTTAGACCCGAGGCTCAACAAGGACTCCAAAAGA---TTAAGGACCTAAAAGCCCAAGG 176

Qy       2591 CTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGA 2650
             | |||||||| ||||||  |  |||| || |||| |        | |   |  | |||||
Db        177 CCTAGTAAAAGCATGCAATAGCCCCTACAATAATCC-------AACTTTAGGAGTACAGA 229

Qy       2651 AACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGT 2710
             |||||||||||||||||| ||||||||||||||||||||||||||||| |||| |  |||
Db        230 AACCCAGTGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGTCACTGT 287

Qy       2711 CCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGC 2770
             || | |||||| |||||||||| |||||||| | ||::|||||||||||||||||||||
Db        288 CCCTCTATACCTAGCTGTACCTAACCCTTATATTCTGCTTTCCCAAATACCAGAGGAAGC 347

Qy       2771 AGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGA 2830
             |||||||||||||: ||||||||||:|||||||||| ||||||||||||||||||||||
Db        348 AGAGTGGTTTACAGACCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGA 407

Qy       2831 CTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGAC 2890
             |||||||||||| |||||||||||||| ||||||||||:  |||||||||||||||||||
Db        408 CTCTCAATTCTTATTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAACTCACCTGGAC 467

Qy       2891 TRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGA 2950
             |:||| |||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        468 TGTTTCACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGA 527

Qy       2951 CTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             ||||||:| :|::|||||||||| ||||||||||||
Db        528 CTTGAGCCGGTTCTCATACCTGGGCACTCTTGTCCTT 564


RESULT 175
ACH73335
ID   ACH73335 standard; DNA; 507 BP.
XX
AC   ACH73335;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #6530.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 6530; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 507 BP; 154 A; 112 C; 115 G; 126 T; 0 U; 0 Other;

  Query Match             5.6%;  Score 423;  DB 1;  Length 507;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  434;  Conservative   40;  Mismatches   21;  Indels    6;  Gaps    2;

Qy       3399 ACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAA 3458
             |||||||:|||||||||||||:|||||||  :|||||||||||||:||:|||::|||||| |
Db         11 ACCCAAGCCCCAGTGTTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATATGTCACAG-A 69

Qy       3459 AAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRT 3518
             ||||||||:||:||||| :|||||||||| |||:|||:|||||:|||| |||||||::|
Db         70 AAAAACAGGAATAGCTCCAGGAGTCCTTACGCAGATCCAAGGGACGAGCCTGCAACCCAT 129

Qy       3519 GGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGT 3578
             |||::||||||:||||||||:     |||||||||||||||||:|:||||||||||:|||| 
Db        130 GGCATACCTGAGTAAGGAAAT-----TAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGT 184
```

```
Qy       3579 AGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCT 3638
Db        185 AGTGGCAGCAGTCACAGTCTTAGTAACTGAAGCAGTTAAAATGATACAAGGAAGAGATCT 244

Qy       3639 TACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCT 3698
Db        245 TACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGACT 304

Qy       3699 GTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACT 3758
Db        305 GTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGTTGCGACT 364

Qy       3759 GTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARA 3818
Db        365 GTGCACTTGTGCAACTCTTAACCCAGCCACATTGCTTCCAGACAATGAAGAAAAGATAGA 424

Qy       3819 AYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGT 3878
Db        425 ACATAACTGTCAACAAATAATTGCTCAAACCTACACTGCTCGAGGGGACCTTTTAGAAGT 484

Qy       3879 TCCYTTGACTGATCCYGACCT 3899
Db        485 TCCCTTGACTGATCCCGATCT 505
```

RESULT 176
AAI44191
ID AAI44191 standard; DNA; 558 BP.
XX
AC AAI44191;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #12877 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 12877; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 558 BP; 181 A; 132 C; 117 G; 128 T; 0 U; 0 Other;

Query Match 5.6%; Score 420.9; DB 1; Length 558;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 512; Conservative 8; Mismatches 36; Indels 11; Gaps 7;

```
Qy        780 GTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACA 839
Db          1 GAAAAGAAGGCAAATGGAGTGAAGTGCCATATGTACAAACTTTCTTTTCATTAAGAGACA 60

Qy        840 ACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTC 899
Db         61 ACTCGCAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTCTCAGAGTCTACCTC 120

Qy        900 CCTATCCCAGCATCCCC--GACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
Db        121 CCTACCCCGGCATCCCCCTGACTCCTTCTCCAACTAATAAGCAACCCCCTTCAACCCAAA 180

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
Db        181 TGGTCCAAAAGGAGATAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCC 240

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
Db        241 GTTTATGCCCCCT-CCAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTA 299

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
Db        300 CCTTTTTCTCTCTCAGACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAAC 359

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
Db        360 CCTGATGGCTATATTGATG-TTTACAAGGGTTAGGACAA---TCTGATCTGACATGGAGA 415

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
Db        416 GATGTA-ATGTTACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 474

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
Db        475 TGCAGCCCGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ACAGGATGACA 532

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCAC 1344
Db        533 ACAG-AGGAAAGAGAACGATTGCCCAC 558
```

RESULT 177
ACH71117/c
ID ACH71117 standard; DNA; 521 BP.
XX
AC ACH71117;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #4312.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX

```
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 4312; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 521 BP; 140 A; 102 C; 128 G; 151 T; 0 U; 0 Other;

  Query Match             5.6%;  Score 420.2;  DB 1;  Length 521;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  456;  Conservative   10;  Mismatches   43;  Indels   12;  Gaps    1;

Qy       6642 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 6701
             |||||| |||||||||| ||||| |||||||||||||||||||||||||||||||||||||
Db        521 AAATGGTGACATGGAATGGGTCACCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 462

Qy       6702 CCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGG 6761
             | |||||||||||||||||:|||||||||||||||||:|||||| || ||:|||||
Db        461 CTTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCCGAAAGAGG 402

Qy       6762 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGT 6821
             ||||| |||||||||||||| |||||||||||||||||| |||||||||| ||||||||
Db        401 GGGAATCTGTTTATTTTTAGTGGAAGAATGCTGTTATTACGTTAATCAATCTGGAATCGT 342

Qy       6822 CACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAA 6881
             ||| || ||||||:||||||||| |||||||||||||| ||:||||||:||||||| |||
Db        341 CACCGAAAAAGTTAAAGAAATTCAAGATCGAATACAACATAGAGCAGAGGAGCTTCAAAA 282

Qy       6882 CACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACC 6941
             |||  ||||||||| ||||||||||:|||||||||||||||||||||||||||||||||||
Db        281 CACCAGACCCTGGGACCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACC 222
```

```
Qy       6942 TCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAA 7001
             ||||||  |||||||| || |||||||||||||||||||||||||:|||||||| ||||
Db        221 TCTAGCGACTATAATACTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTTTTAA 162

Qy       7002 CTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAA 7049
             |||||||||||||||| |||||||:|| |||           |||||||||||| |
Db        161 GTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTACAAATGGCTCTTCAAATGGAGCCCCA 102

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||| || |||| |||||| |||||| |  |||||||||||| |||||:|||  || | ||
Db        101 GATACAATCCATGACTAAAATCTACTGTGGACCCCTGGACCAGCCTGCTAGTGCATGCTC 42

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCT 7150
              |||||||||||||  ||||||||||||||||||||||||
Db         41 CCATGTTAATGACATTGAAGGCACCCCTCCTGAGGAAATCT 1
```

```
RESULT 178
AAI11856
ID   AAI11856 standard; DNA; 488 BP.
XX
AC   AAI11856;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #1789 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 1789; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;
```

```
Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | |||||||||  |||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||  |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             |||||  ||||||||  |||||||  ||||||||||  ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             ||||  ||  |  |||||||||||||||||||||||||  |  |||||  ||||||||||  |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 179
ABA53554
ID ABA53554 standard; DNA; 488 BP.
XX
AC ABA53554;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #1859.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 1; SEQ ID NO 1859; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

Query Match 5.5%; Score 409.6; DB 1; Length 488;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 451; Conservative 1; Mismatches 35; Indels 2; Gaps 2;

```
Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | |||||||||  |||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||  |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             |||||  ||||||||  |||||||  ||||||||||  ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             ||||  ||  |  |||||||||||||||||||||||||  |  |||||  ||||||||||  |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 180
AAI33179
ID AAI33179 standard; DNA; 488 BP.
XX
AC AAI33179;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #1865 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX

```
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 1865; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             |||||||||||  |||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||  | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488

RESULT 181
ABA43137
ID   ABA43137 standard; DNA; 488 BP.
XX
AC   ABA43137;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #1832.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 1832; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             |||||||||||  |||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120
```

```
Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||| |||||| |||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||| ||||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||||  | ||||| |||||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 182
ABA23311
ID ABA23311 standard; DNA; 488 BP.
XX
AC ABA23311;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #1777 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 1; SEQ ID NO 1777; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

```
  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             |||||||||||  |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||| ||||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||||  | ||||| |||||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 183
AAK27280
ID AAK27280 standard; DNA; 488 BP.
XX
AC AAK27280;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 1837.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX

```
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 1837; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||| |||||| |||||||||||||||||||| || |||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||| | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488


RESULT 184
AAK01823
ID   AAK01823 standard; DNA; 488 BP.
XX
AC   AAK01823;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 1814.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 1814; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:|||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||| |||||| |||||||||||||||||||| || |||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||||
```

```
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | ||||||||||||||||||||||||  | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||| |||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 185
ABS26854
ID ABS26854 standard; DNA; 488 BP.
XX
AC ABS26854;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 1844.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 1; SEQ ID NO 1844; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

```
  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
              |||||||||||| |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
              | ||||||||| |||||||||||||||||||||| ||||||| |||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
              |||||||||||||||:||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
              |||||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
              ||||| |||||||| ||||||| |||||||||||  ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | ||||||||||||||||||||||||  | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||| |||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488
```

RESULT 186
AAI01794
ID AAI01794 standard; DNA; 488 BP.
XX
AC AAI01794;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #1785 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

```
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 1785; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             |||||||||||| |||||||||||||||||||||||   ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | |||||||| |||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             |||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||| |||||| |||||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||||||  | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488


RESULT 187
ABS01806
ID   ABS01806 standard; DNA; 488 BP.
XX
AC   ABS01806;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 1797.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 1797; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
```

```
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 488 BP; 155 A; 107 C; 116 G; 110 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | |||||||||  |||||||||||||||||||||| |||||||  |||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             ||||||||||||||||:||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||  |||||| ||||||||||||||||||||  || |||||||||  ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             |||||  ||||||||  |||||||  ||||||||||  ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCG 1935
             ||||  ||  |  |||||||||||||||||||||||||  |  |||||  ||||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||| ||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 188
ABA60281/c
ID   ABA60281 standard; DNA; 504 BP.
XX
AC   ABA60281;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #8586.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 8586; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAA 7038
             |||||| |||| || ||||||||||||||||||||| ||:|| |||           |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||   ||| |||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||||  |||||||||||||||| || |||||||||||| ||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db         84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 189
AAI40167/c
ID   AAI40167 standard; DNA; 504 BP.
XX
AC   AAI40167;
XX
DT   17-OCT-2001  (first entry)
XX
```

```
DE   Probe #8853 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 8853; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||            |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||    ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
               ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||| |||||||||||||||
Db         84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 190
ABA28561/c
ID   ABA28561 standard; DNA; 504 BP.
XX
AC   ABA28561;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #7027 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 7027; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:||
```

```
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAA 7038
              |||||| |||| || ||||||||||||||||||||| ||:|| |||            |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
              ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
              |||||| | |||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 191
AAK34442/c
ID AAK34442 standard; DNA; 504 BP.
XX
AC AAK34442;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 8999.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 8999; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

```
  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
              ||||||| || |||  |||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
              |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAA 7038
              |||||| |||| || ||||||||||||||||||||| ||:|| |||            |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
              ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
              |||||| | |||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 192
AAK08559/c
ID AAK08559 standard; DNA; 504 BP.
XX
AC AAK08559;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 8550.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.

```
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 8550; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||: |||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||          |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| ||||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
               ||| |||||| |||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             |||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 193
ABS34221/c
ID   ABS34221 standard; DNA; 504 BP.
XX
AC   ABS34221;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 9211.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 9211; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||: |||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||          |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| ||||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
```

```
              |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              |||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 194
ABS09114/c
ID ABS09114 standard; DNA; 504 BP.
XX
AC ABS09114;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 9105.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 1; SEQ ID NO 9105; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 504 BP; 117 A; 108 C; 150 G; 129 T; 0 U; 0 Other;

Query Match 5.3%; Score 400.4; DB 1; Length 504;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 438; Conservative 9; Mismatches 45; Indels 12; Gaps 1;

```
Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
              ||||||| || |||  |||||||||||||||||||||: |||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
              |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
              |||||| |||| || |||||||||||||||||||| ||:|| |||          |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
              ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
              |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              |||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
```

```
Db         24 CACGGGGCTTGCAACTTAGCTCAC 1

RESULT 195
ACH90788
ID   ACH90788 standard; DNA; 474 BP.
XX
AC   ACH90788;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #23983.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 23983; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 474 BP; 155 A; 96 C; 108 G; 115 T; 0 U; 0 Other;

  Query Match             5.2%;  Score 387.5;  DB 1;  Length 474;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  403;  Conservative   40;  Mismatches   31;  Indels    3;  Gaps    2;

Qy       3344 TAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCA 3403
             |||||||||:|:||||||:::||| |||||||||||||||  ||  | | |||  ||||||
Db          1 TAAGATGGACACCTGAAGCAAAAGAGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCA 58

Qy       3404 AGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAA 3463
             ||:|||| |||||||:||||||||:||||||| ||| |:||:|||::|||||| ||||||
Db         59 AGCCCCAATGTTAAGTTTGCCAACAGGGCAAGGCTTCTCTTTATATGTCACAG-AAAAAA 117

Qy       3464 CAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRY 3523
             |||:||:||||||:||| |||||||||||:| |:|||||: ||||||||||::||||::
Db        118 CAGGAATAGCTCTAGGAATCCTTACACAGGTTCAAGGGACCAGCTTGCAACCTGTGGCAT 177

Qy       3524 ACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGG 3583
             || |||:|||| |||:|||||||||||||||||||:|:||||||||: |||||||
Db        178 ACTTGAGTAAGAAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGA 237

Qy       3584 TGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTG 3643
             ||||||||||||::||||||||||||||||||||||||||||||:|||||||||||||||
Db        238 CGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTG 297

Qy       3644 TGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAG 3703
             |||||||||||||||||:|| ||::| |||||||||||||||||| ||||||||||||||||
Db        298 TGTGGACATCTCATGATGTAAATGGTATACTCACTGCTAAAGGAAACTTGTGGCTGTCAG 357

Qy       3704 ACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCA 3763
             |||||: ||| |||||| :|||||||||||||| |||||:| | ||||||| :|||| |||
Db        358 ACAACCATTTGCTTAAAAATCAGGCTCTATTAATTGAAGGACTAGTGCTGAGACTGCGCA 417

Qy       3764 CTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAY 3820
             ||||||||||||||||||||:| |||||||||||||||||||||||||||||| |:||:
Db        418 CTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGAGAGAAC 474


RESULT 196
ADG14942
ID   ADG14942 standard; cDNA; 438 BP.
XX
AC   ADG14942;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV associated cDNA #27.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
```

```
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Disclosure; SEQ ID NO 201; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 438 BP; 118 A; 104 C; 68 G; 92 T; 0 U; 56 Other;

  Query Match             5.1%;  Score 382.5;  DB 1;  Length 438;
  Best Local Similarity   79.9%;  Pred. No. 0;
  Matches  350;  Conservative   82;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||:|:||||||:::|||||||:|||:||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              :|:||: ||::|||||||||||||||:||:||||||||||||:||||||:||||||||:
Db         61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:::|||||:||||||||||:|||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||:|:|||:::|||||||:||||||:::|||||||:|
Db        181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCST 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:||||||||||||:::::|||||||||:||::|
Db        241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:|||||||:||:|||||| |||:|:|:  |:||::|||||||::||||||:||||::
Db        301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::||||||||||||:|||||||:|||||||:||:|||||||||:|:||||:::::||||
Db        360 ADYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 197
AAV43239
ID   AAV43239 standard; cDNA; 438 BP.
XX
AC   AAV43239;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 43 pol gene region from patient 68-1.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .438
FT                   /*tag=  a
FT                   /product= "Protein AAW71075"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71075.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 201; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 43 obtained from an MS
CC   patient 68-1. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 438 BP; 138 A; 116 C; 78 G; 104 T; 0 U; 2 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||:|||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

```
RESULT 198
AAV43237
ID   AAV43237 standard; cDNA; 438 BP.
XX
AC   AAV43237;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 41 pol gene region from patient 68-1.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .438
FT                   /*tag= a
FT                   /product= "Protein AAW71073"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71073.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 200; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 41 obtained from an MS
CC   patient 68-1. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 438 BP; 137 A; 119 C; 79 G; 103 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:| |||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||| ||||||||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 199
ADB84424
ID   ADB84424 standard; DNA; 438 BP.
XX
AC   ADB84424;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region DNA from patient sample #4.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84429.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 50A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
```

```
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 438 BP; 137 A; 119 C; 79 G; 103 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  |||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:||||||| |||||||:  ||:| |||||||| :|||||:||||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              ||||||||||||||| |||||||||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 200
ADB84426
ID   ADB84426 standard; DNA; 438 BP.
XX
AC   ADB84426;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region DNA from patient sample #5.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84428.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 50A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 438 BP; 138 A; 116 C; 78 G; 104 T; 0 U; 2 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:|||||||||||:|||  |||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:||||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180
```

```
Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
             ||||||||:||||:||||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
             ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
             ||:||||||| ||:||||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
             |||||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
             ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 201
ADB84423
ID ADB84423 standard; DNA; 438 BP.
XX
AC ADB84423;
XX
DT 04-DEC-2003 (first entry)
XX
DE MSRV-1 pol region consensus #2 DNA from patient sample.
XX
KW MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003039664-A1.
XX
PD 27-FEB-2003.
XX
PF 26-NOV-1997; 97US-00979847.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
DR P-PSDB; ADB84497.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Example 19; Fig 50A; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC The SEQ ID numbers for the sequences as displayed in the main body of the
CC patent do not match the SEQ ID numbers in the sequence listing.
CC Consequently those sequences mentioned in the claims may not be the
CC sequences the authors intended to claim.
XX
SQ Sequence 438 BP; 135 A; 116 C; 78 G; 103 T; 0 U; 6 Other;

Query Match 5.1%; Score 381.3; DB 1; Length 438;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 375; Conservative 45; Mismatches 17; Indels 1; Gaps 1;

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||:|||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
             ||:||: ||::||||||||||||||||||||||||||||||||:|||||:||||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
             ||:||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
             ||||||||:||||:||||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
             ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
             ||:||||||| ||:||||||| |||||||:  ||:|::||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
             |||||||||||||||| |||||||:||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
             ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 202
ADG14872
ID ADG14872 standard; cDNA; 438 BP.
XX
AC ADG14872;
XX
DT 26-FEB-2004 (first entry)
XX
DE MSRV cDNA #4.
XX
KW ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS Multiple sclerosis associated retrovirus.
XX

```
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 131; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 438 BP; 138 A; 116 C; 78 G; 104 T; 0 U; 2 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:||||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 203
ADG14870
ID   ADG14870 standard; cDNA; 438 BP.
XX
AC   ADG14870;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #2.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 129; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 438 BP; 137 A; 119 C; 79 G; 103 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:| |||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:|:|||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
```

```
                    ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 204
ADG14869
ID   ADG14869 standard; cDNA; 438 BP.
XX
AC   ADG14869;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #1.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 128; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 438 BP; 135 A; 116 C; 78 G; 103 T; 0 U; 6 Other;

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  375;  Conservative   45;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||||| ||:|||||||| |||||||:  ||:|::|||||||| :|||||:|||||:
Db         301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| |||||||:||||||||||:|||||||||:|:|||||||:::||||
Db         360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db         420 GGCTTTCCAGGCCCTAAA 437


RESULT 205
AAV43238
ID   AAV43238 standard; cDNA; 438 BP.
XX
AC   AAV43238;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 42 pol gene region from patient 68-1.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .438
FT                   /*tag=  a
FT                   /product= "Protein AAW71074"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71074.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 201; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 42 obtained from an MS
CC   patient 68-1. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 438 BP; 136 A; 119 C; 80 G; 103 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;
```

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||| |||||||||:|||||||||:|:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 206
ADB84425
ID ADB84425 standard; DNA; 438 BP.
XX
AC ADB84425;
XX
DT 04-DEC-2003 (first entry)
XX
DE MSRV-1 pol region DNA from patient sample #6.
XX
KW MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003039664-A1.
XX
PD 27-FEB-2003.
XX
PF 26-NOV-1997; 97US-00979847.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2003-512253/48.
DR P-PSDB; ADB84427.
XX
PT New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS Example 19; Fig 50A; 193pp; English.
XX
CC The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs).Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC The SEQ ID numbers for the sequences as displayed in the main body of the
CC patent do not match the SEQ ID numbers in the sequence listing.
CC Consequently those sequences mentioned in the claims may not be the
CC sequences the authors intended to claim.
XX
SQ Sequence 438 BP; 136 A; 119 C; 80 G; 103 T; 0 U; 0 Other;

Query Match 5.1%; Score 379.7; DB 1; Length 438;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 377; Conservative 41; Mismatches 19; Indels 1; Gaps 1;

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||| |||||||||:|||||||||:|:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 207
ADG14871

```
ID   ADG14871 standard; cDNA; 438 BP.
XX
AC   ADG14871;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #3.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 130; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 438 BP; 136 A; 119 C; 80 G; 103 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|:|||||:||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||| ||||||| ||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 208
AAI36615
ID   AAI36615 standard; DNA; 433 BP.
XX
AC   AAI36615;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #5301 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 5301; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 433 BP; 141 A; 90 C; 103 G; 99 T; 0 U; 0 Other;

  Query Match             5.1%;  Score 378.6;  DB 1;  Length 433;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  377;  Conservative   38;  Mismatches   17;  Indels    2;  Gaps    1;

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||:||||:|||||||||||:|||||:||||  :|||||||||||||||||||||||
Db          2 CCCAGGTATGGCGAAATAGCCAGGCCATTATATACGCTAATTAAGGAAACTCAGAAAGCC 61

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:|:|| |||:::|||||||||||||||||||  |    |
Db         62 AATACCCATTTAGTAAGATGGACACCTAAAGCAGAAGTGGCTTTCCAGGCCCTAACGAAG 121

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:|||||| |||||:||||||||:||||||||||||:||:|||::
Db        122 GCCCT--AACCCAAGCCCCAGTATTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATG 179

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              ||||||||||||||||:||:| ||||:|||||||||||||:|||:||| |:||||||||
```

```
Db        180 TCACAGAAAAAAACAGGAATAACTCTAGGAGTCCTTACACAGGTCCAAGGAACGAGCTTG 239

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
              |||||::||||: ||||||:||||||||:|||||||||||||||||||||:|:|||||||
Db        240 CAACCCGTGGCAGACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 299

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
              |||:|||||| |||||||||||||| |::|||||||||||||||||||||||||||||| :
Db        300 TTATGGGTAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTTAAAATAATACAGGAA 359

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
              |||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||||
Db        360 AGAGATCTTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGAC 419

Qy       3691 TTGTGGCTGTCAGA 3704
              ||||||||||||||
Db        420 CTGTGGCTGTCAGA 433
```

RESULT 209
AAI13992
ID AAI13992 standard; DNA; 422 BP.
XX
AC AAI13992;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #3925 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 3925; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 422 BP; 122 A; 83 C; 100 G; 117 T; 0 U; 0 Other;

Query Match 5.0%; Score 377.7; DB 1; Length 422;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 384; Conservative 22; Mismatches 16; Indels 1; Gaps 1;

```
Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db          1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |  |||||||| ||||||||:|||||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              ||||||||||||:|||||||||||:|||| ||||||||| |||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              |||||||||||||||||||||||||||:||:||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
              |||||| |||||||||| ||:||||:||||||:|||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
              |||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
              |||
Db        420 TGA 422
```

RESULT 210
AAI35373
ID AAI35373 standard; DNA; 422 BP.
XX
AC AAI35373;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4059 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.

```
XX
PS   Claim 25; SEQ ID NO 4059; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 422 BP; 122 A; 83 C; 100 G; 117 T; 0 U; 0 Other;

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
             ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
             |  |||||||| ||||||||:||||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
             ||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
             ||||||||||||:|||||||||||:|||| |||||||| |||||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
             ||||||||||||||||||||||||||:||:||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
             |||||| |||||||||||| ||:||||:|||||:||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             ||||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
             |||
Db        420 TGA 422


RESULT 211
ABA45226
ID   ABA45226 standard; DNA; 422 BP.
XX
AC   ABA45226;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3921.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 3921; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 422 BP; 122 A; 83 C; 100 G; 117 T; 0 U; 0 Other;

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
             ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
             |  |||||||| ||||||||:||||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
             ||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
             ||||||||||||:|||||||||||:|||| |||||||| |||||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
             ||||||||||||||||||||||||||:||:||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
             |||||| |||||||||||| ||:||||:|||||:||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             ||||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
             |||
Db        420 TGA 422
```

```
RESULT 212
AAK29413
ID   AAK29413 standard; DNA; 422 BP.
XX
AC   AAK29413;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3970.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 3970; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 422 BP; 122 A; 83 C; 100 G; 117 T; 0 U; 0 Other;

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
             ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db          1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
             |  |||||||| ||||||||:||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
             ||||||||||||||||||||||||||||||||||:||||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
             ||||||||||||:|||||||||||||:|||| |||||||||| ||||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
             |||||||||||||||||||||||||:||:||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
             |||||| |||||||||||  ||:||||:|||||:||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             |||||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
             |||
Db        420 TGA 422


RESULT 213
ABS29038
ID   ABS29038 standard; DNA; 422 BP.
XX
AC   ABS29038;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 4028.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 4028; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 422 BP; 122 A; 83 C; 100 G; 117 T; 0 U; 0 Other;

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
```

```
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||| |||:|:||||||||||: |||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |  |||||||| ||||||||:||||||||||||||||||||||||||:|||||:: |
Db            61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              ||||||||||||||||||||||||||||||||||||||:|||||||||||:|||||||
Db           121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              ||||||||||:||||||||||||:|||| |||||||| ||||||||||||||:| ||
Db           181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              ||||||||||||||||||||||||:||:||||||| ||||||:|||| ||||||||
Db           241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
              |||||| |||||||||||| ||:||||:|||||:||||||||||: ||| ||||||||||
Db           301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
              |||||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||
Db           360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
              |||
Db           420 TGA 422


RESULT 214
ADB84496
ID   ADB84496 standard; DNA; 438 BP.
XX
AC   ADB84496;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #43.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Disclosure; Page 112; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 438 BP; 115 A; 104 C; 67 G; 91 T; 0 U; 61 Other;

  Query Match             5.0%;  Score 376.9;  DB 1;  Length 438;
  Best Local Similarity   78.8%;  Pred. No. 0;
  Matches  345;  Conservative   81;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||:|:|||||||:::|||||||:|||:||
Db             1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              :|:||: ||::||||||||||||||:||:||||||||||||:|||||||:||||||||
Db            61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTN 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:::|||||:|||||||||||:||||||||
Db           121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAN 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||:|:|||:::|||||||:|||||:::|||||||:
Db           181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCSN 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:||||||||||||:::::|||||||||:||::
Db           241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCN 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:||||||||:|||:|||||||  |||:|:|:  |:||::|||||||::|||||:||||::
Db           301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              ::||||||||||||:|||||||:|||||||:||:||||||||||:|:||||:::::||||
Db           360 NDYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db           420 NGCTTTCCAGGCCCTAAA 437


RESULT 215
```

```
AAS87567/c
ID   AAS87567 standard; cDNA; 448 BP.
XX
AC   AAS87567;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23371.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23380.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23371; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 448 BP; 96 A; 103 C; 117 G; 132 T; 0 U; 0 Other;

  Query Match             4.8%;  Score 357.2;  DB 1;  Length 448;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  383;  Conservative    9;  Mismatches   27;  Indels   10;  Gaps    1;

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
             |||||||||| |||| | |||||| ||||||||||||||| |||||||||:||| | |||
Db        448 CTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCAGTCAA 389

Qy       7214 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 7273
             |||||||| |||||||| |||||||||||||||| |||| ||||||||||||||||||||
Db        388 CCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCT 329

Qy       7274 GGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCT 7333
             |||||||||||||| ||:|||||||||:|||||||||:|||||||||||||| | |||||||
Db        328 GGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCT 269

Qy       7334 TTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 7393
             |||||||||||||||||||||||||:||||||| ||||||||||||||||||||||||||||
Db        268 TTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGC 209

Qy       7394 TAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGC 7453
             |||||| ||||| |||||||||||||||||||||| ||||:||| ||:||||||||||||
Db        208 TAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGC 149

Qy       7454 AGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTT 7513
             | |||||||||||||||||||||||||||||| |  |||||||| |||||||| |||| |||
Db        148 AAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTT 89

Qy       7514 TGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTT 7573
             |||||||||||||||||||||||||| ||||          ||||||||||||||||||
Db         88 TGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTCACTCTATTAAATCTT 39

Qy       7574 GCARCTGCR 7582
             |||:||||:
Db         38 GCAACTGCA 30


RESULT 216
AAI63993/c
ID   AAI63993 standard; DNA; 439 BP.
XX
AC   AAI63993;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 365.
XX
KW   Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW   neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
KW   antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW   cardiovascular disorder; neurological disease; infection; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155308-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001309.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
```

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488781/53.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders.
XX
PS Disclosure; SEQ ID NO 365; 664pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI63803-AAI64012) and
CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 439 BP; 99 A; 95 C; 114 G; 131 T; 0 U; 0 Other;

```
  Query Match             4.7%;  Score 353.6;  DB 1;  Length 439;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   32;  Indels    6;  Gaps    2;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||| || ||| |||||| |||||||||||||||||||||||| ||:
Db        439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||| ||||||| |||||||||||||||  ||||||||||    ||
Db        379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:|||||||||:|||||||||:||||||||| || ||
Db        323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             |||||  ||||||||| ||||||||||||||||:|||||||  ||||||||||||||||||
Db        263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||| |||||||||||||||||||||||||:||||||:||||
Db        203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| ||||||||| ||| |  ||| |||| |||| ||||
Db        143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             |||||||||||||| |||||  ||||||||||||||||||||  ||||||||||||||
Db          83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db          25 TAAATCTTGCACCTGCA 9
```

RESULT 217
AAS31671/c
ID AAS31671 standard; DNA; 439 BP.
XX
AC AAS31671;
XX
DT 04-DEC-2001 (first entry)
XX
DE Genomic sequence #17 encoding novel human calcium-binding protein.
XX
KW Human; calcium-binding protein; calcium flux; neurological disease;
KW immune dysfunction; digestive disorder; neoplastic disease;
KW blood disorder; infectious disease; gene therapy; immunosuppressive;
KW antiarthritic; cytostatic; vasotropic; antibacterial; nootropic;
KW virucide; ds.
XX
OS Homo sapiens.
XX
PN WO200155304-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001302.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.

```
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465568/50.
XX
PT   Isolated nucleic acid molecule encoding a calcium-binding protein is used
PT   in preventing, treating or ameliorating a medical condition.
XX
PS   Disclosure; SEQ ID NO 183; 542pp; English.
XX
CC   The present invention relates to the isolation of novel human calcium-
CC   binding proteins (AAU19892-AAU19969), and cDNA and genomic sequences
CC   encoding for these proteins. The sequences of the invention are useful in
CC   the diagnosis, prevention and/or prognosis of diseases associated with
CC   aberrant calcium flux. Such disorders include neurological diseases (e.g.
CC   amylotrophic lateral sclerosis, ALS), immune dysfunction (e.g. severe
CC   combined immunodeficiency, SCID), digestive disorders (e.g. irritable
CC   bowel syndrome, IBS), neoplastic disease (e.g. cancer), blood disorders
CC   (e.g. haemophilia), and/or infectious disease (e.g. acquired
CC   immunodeficiency syndrome, AIDS). The novel calcium-binding proteins are
CC   also useful as screening tools to identify antagonists and/or agonists
CC   that may enhance or inhibit activities mediated by calcium-binding
CC   proteins. The polynucleotides of the invention are also useful in gene
CC   therapy. AAS31655-AAS31680 represent genomic sequences encoding for the
CC   novel human calcium-binding proteins. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 439 BP; 99 A; 95 C; 114 G; 131 T; 0 U; 0 Other;

  Query Match             4.7%;  Score 353.6;  DB 1;  Length 439;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   32;  Indels    6;  Gaps    2;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||| || ||| |||||| |||||||||||||||||||||||| ||:
Db        439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||| ||||||| |||||||||||||||  |||||||||    ||
Db        379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:|||||||:|||||||:|||||||| || ||
Db        323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             ||||| |||||||| |||||||||||||||:|||||| |||||||||||||||||
Db        263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||||| ||||||||||||||||||||||||||:||||||:||||
Db        203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| |||||||| ||| |  ||| |||| |||| ||||
Db        143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             ||||||||||||| |||||  |||||||||||||||||||  ||||||||||||||
Db          83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db          25 TAAATCTTGCACCTGCA 9


RESULT 218
ADM24708/c
ID   ADM24708 standard; cDNA; 439 BP.
XX
AC   ADM24708;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human PRO polynucleotide #191.
XX
KW   Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW   cardiovascular disorder; renal disorder; proliferative disorder; cancers;
KW   systemic lupus erythematosus; rheumatoid arthritis; multiple sclerosis;
KW   thyroiditis; anaemia; Grave's disease; diabetes; hepatitis; asthma;
KW   allergy; nephritis; Parkinson's disease; Alzheimer's disease;
KW   atherosclerosis; myocardial infarction; AIDS; infection.
XX
OS   Homo sapiens.
XX
PN   US2004014039-A1.
XX
PD   22-JAN-2004.
XX
```

PF 31-MAY-2002; 2002US-00158057.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764890.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2004-108205/11.
XX

```
PT   Secreted and transmembrane PRO polypeptides and nucleic acids for
PT   diagnosing, preventing or treating diseases associated with aberrant
PT   expression or activity of the polypeptide, e.g. cancer, asthma, AIDS,
PT   Parkinson's disease or diabetes.
XX
PS   Claim 4; SEQ ID NO 365; 284pp; English.
XX
CC   The invention relates to a secreted and transmembrane PRO polypeptides
CC   and the polynucleotides encoding them. The polypeptides and
CC   polynucleotides are useful in diagnosing, preventing, prognosing or
CC   treating diseases or disorders associated with aberrant expression and/or
CC   activity of PRO polypeptides, such as neural disorders, immune system
CC   disorders, muscular disorders, reproductive disorders, gastrointestinal
CC   disorders, pulmonary disorders, cardiovascular disorders, renal
CC   disorders, proliferative disorders and/or cancers. In particular, these
CC   diseases are systemic lupus erythematosus, rheumatoid arthritis, multiple
CC   sclerosis, thyroiditis, anaemia, Grave's disease, diabetes, hepatitis,
CC   asthma, allergies, nephritis, Parkinson's disease, Alzheimer's disease,
CC   atherosclerosis, myocardial infarction, AIDS, infections, etc. This
CC   sequence represents a human PRO polynucleotide of the invention. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 439 BP; 99 A; 95 C; 114 G; 131 T; 0 U; 0 Other;

  Query Match             4.7%;  Score 353.6;  DB 1;  Length 439;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   32;  Indels    6;  Gaps    2;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||| || ||| |||||| |||||||||||||||||||||||| ||:
Db        439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||||| |||||||| |||||||||||||||  ||||||||||    ||
Db        379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:||||||||:||||||||:||||||||| || ||
Db        323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             ||||| |||||||||| |||||||||||||||:|||||||  |||||||||||||||||
Db        263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||| |||||||||||||||||||||||:||||||:||||
Db        203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| ||||||||| ||| |  ||| |||| |||| ||||
Db        143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             |||||||||||||| |||||  |||||||||||||||||||  ||||||||||||||||
Db         83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db         25 TAAATCTTGCACCTGCA 9


RESULT 219
AAV86331
ID   AAV86331 standard; cDNA; 384 BP.
XX
AC   AAV86331;
XX
DT   27-APR-1999  (first entry)
XX
DE   EST clone AJ172.
XX
KW   Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW   tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW   chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW   receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS   Homo sapiens.
XX
PN   WO9845435-A2.
XX
PD   15-OCT-1998.
XX
PF   10-APR-1998;   98WO-US006954.
XX
PR   10-APR-1997;   97US-00835913.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Racie LA,  Merberg D,  Treacy M;
PI   Spaulding V,  Agostino MJ;
XX
DR   WPI; 1999-070076/06.
XX
PT   New polynucleotides encoding human secreted proteins - derived from e.g.
PT   human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT   pituitary, retina and colon cDNA libraries.
XX
PS   Claim 1; Page 199; 633pp; English.
XX
CC   This sequence represents an expressed sequence tag (EST), and is a
CC   polynucleotide of the invention. The polynucleotides of the invention are
CC   all secreted EST sequences isolated from a variety of human tissue
CC   sources. The EST sequences and proteins encoded by them are predicted to
CC   have useful biological activities which would make them suitable for
CC   treating, preventing or ameliorating medical conditions in humans and
CC   animals, although no supporting data is given. Suggested activities
CC   include nutritional activity, immune stimulating or suppressing activity,
CC   haematopoiesis regulating activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC   activity. The EST sequences are also stated to be useful for gene therapy
XX
SQ   Sequence 384 BP; 100 A; 120 C; 65 G; 99 T; 0 U; 0 Other;

  Query Match             4.7%;  Score 352.9;  DB 1;  Length 384;
  Best Local Similarity   96.0%;  Pred. No. 0;
  Matches  361;  Conservative    9;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCTGCCTCGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 60

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 120

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        121 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 179

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             |||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||
Db        180 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 239

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             |||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        240 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 299

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        300 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 359

Qy       5772 CCACACCCATATGCCC 5787
```

```
                ||||||||||||||||
Db        360 CCACACCCATATGCCC 375


RESULT 220
AAV43243
ID   AAV43243 standard; cDNA; 429 BP.
XX
AC   AAV43243;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 8 pol gene region from patient 46-7.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .429
FT                   /*tag= a
FT                   /product= "Protein AAW71078"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71078.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 204; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 8 obtained from an MS
CC   patient 46-7. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 429 BP; 132 A; 116 C; 82 G; 98 T; 0 U; 1 Other;

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| || |||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:|||||:|||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |||||:||||||| |||||:::|||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:||||||||||||||||||::::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||||| |:  ||:||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  |||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||| ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 221
ADB84430
ID   ADB84430 standard; DNA; 429 BP.
XX
AC   ADB84430;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region DNA from patient sample #1.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84434.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 49A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
```

```
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 429 BP; 126 A; 108 C; 79 G; 96 T; 0 U; 20 Other;

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   79.5%;  Pred. No. 0;
  Matches  348;  Conservative   59;  Mismatches   21;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| |:||:|||:: ||||||:||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCYTCAGTATRGGGATGAYTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| ||:|||||||||||||::||||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          |||||| |:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAARGGCTCASCTCTGCTCACASCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |:|||:|||||||| |||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:|||||||||||||:::||||||||||:||::|
Db        232 ATACTGGYTTATCCYCATCCCAWAACCYTAAAGCAACTAAGARGGTTCCTTGGCATAWCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||||||: | |:  |::|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCV-GATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||||||||||||||||||| ||:|||||||||:|:||||: :::||||
Db        351 ADYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGARACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 222
ADG14879
ID   ADG14879 standard; cDNA; 429 BP.
XX
AC   ADG14879;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #8.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 138; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 429 BP; 132 A; 116 C; 82 G; 98 T; 0 U; 1 Other;

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| || |||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:|||||:||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |||||:|||||||| |||||||:::|||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||::::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||||| |:  ||:|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  ||||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 223
ADB84433
```

```
ID   ADB84433 standard; DNA; 429 BP.
XX
AC   ADB84433;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region consensus #1 DNA from patient sample.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84494.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 49A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 429 BP; 132 A; 115 C; 82 G; 98 T; 0 U; 2 Other;

  Query Match             4.6%;  Score 344.6;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;

  Matches  360;  Conservative   42;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| || |||||||||||||:|||||:||||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |||||:|||||||| |||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||||:::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:|||||||  ||||| |:  ||:||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  ||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 224
AAV43241
ID   AAV43241 standard; cDNA; 429 BP.
XX
AC   AAV43241;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 1 pol gene region from patient 46-7.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .429
FT                   /*tag=  a
FT                   /product= "Protein AAW71076"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71076.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
```

```
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 203; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 1 obtained from an MS
CC   patient 46-7. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 429 BP; 130 A; 113 C; 82 G; 104 T; 0 U; 0 Other;

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| |||||||||||||||||:|||||:||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:|:||||| |||||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:||||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:|| ||||||||||||||||||:|::||||||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| ||| | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||| ||:||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 225
ADB84431
ID   ADB84431 standard; DNA; 429 BP.
XX
AC   ADB84431;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region DNA from patient sample #2.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84435.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 19; Fig 49A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 429 BP; 130 A; 113 C; 82 G; 104 T; 0 U; 0 Other;

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| |||||||||||||||||:|||||:||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:|:||||| |||||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
```

```
           ||||||||:||||:|||||||||||| |||||:||||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
             |||||||||||||||||:|| |||||||||||||||||:|::|||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
             | :||||||||||:|||||| ||| | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
             |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
             ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 226
ADG14877
ID   ADG14877 standard; cDNA; 429 BP.
XX
AC   ADG14877;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #6.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 136; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 429 BP; 130 A; 113 C; 82 G; 104 T; 0 U; 0 Other;

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||| |||||| |||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| |||||||||||||||||:|||||:||||||||
Db          61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:|||||  |||||||||||  |||||||||
Db         121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:|||||||| |||||||:::|||||||
Db         172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||||||:|| |||||||||||||||||:|::|||||||:|| :|
Db         232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||| | |:  | :||||||||||:|||||: |||| |
Db         292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db         351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db         411 GGCTTTCCAGGCCCTAAA 428


RESULT 227
ADG14878
ID   ADG14878 standard; cDNA; 429 BP.
XX
AC   ADG14878;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #7.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 137; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 429 BP; 131 A; 111 C; 85 G; 102 T; 0 U; 0 Other;
```

```
  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  359;  Conservative   39;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
                |:||: ||::|||||||||||||||| ||||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:|:||||| ||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| | |||:|||||||| ||||||:::|||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||| |||||||||||||:|::||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||| | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              | |||||||||||||||||||||||||||||| ||:|||||||||:|:|||| :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428
```

```
RESULT 228
ADG14876
ID   ADG14876 standard; cDNA; 429 BP.
XX
AC   ADG14876;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV cDNA #5.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 19; SEQ ID NO 135; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 429 BP; 127 A; 107 C; 79 G; 97 T; 0 U; 19 Other;
```

```
  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches  351;  Conservative   50;  Mismatches   27;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||:||: ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGAYTTRATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
                |:||: ||::|||||||||||||||| ||:|||||||||||||::|||||:|||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         |||||||:|||||||| ||||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAARGCTCAYCTCTGCTCACAYCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |:|||:|||||||| ||||||:::|||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||:||||:|| |||| |||||||||||||:::|||||||||:|| :|
Db        232 ATACTGGRTTATCCYCATCCCAYAACCRTAAAGCAACTAAGARGGTTCCTTGGCATAYCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| |||:| |:  |::||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCRGATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 ARYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428
```

```
RESULT 229
ABK64210/c
ID   ABK64210 standard; DNA; 410 BP.
XX
AC   ABK64210;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human benign prostatic hyperplasia gene #105.
XX
KW   Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200212440-A2.
XX
PD   14-FEB-2002.
XX
PF   07-AUG-2001; 2001WO-US024708.
XX
PR   07-AUG-2000; 2000US-0223323P.
PR   05-JUN-2001; 2001US-00873319.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (NISB ) JAPAN TOBACCO INC.
XX
```

```
PI   Munger WE,  Kulkarni P,  Getzenberg RH,  Waga I,  Yamamoto J;
XX
DR   WPI; 2002-257476/30.
XX
PT   Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT   detecting expression levels of one or more genes in prostate cells from
PT   patient that are differentially regulated compared to normal prostate
PT   cells.
XX
PS   Disclosure; Page 111; 444pp; English.
XX
CC   The invention relates to a method of diagnosing (I) the onset or
CC   progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC   or identifying an agent that modulates the onset or progression of BPH.
CC   The method is based on changes in gene expression in BPH tissue isolated
CC   from patients exhibiting different clinical states of prostate
CC   hyperplasia as compared to normal prostate tissue. (I) comprises
CC   detecting the expression levels of one or more genes in prostate cells
CC   from the subject that are differentially regulated compared to normal
CC   prostate cells. (II) comprises preparing a first gene expression profile
CC   of BPH cells or BPH-like cell population, exposing the cells to the
CC   agent, preparing a second gene expression profile of the agent exposed
CC   cells, and comparing the first and second gene expression profiles. (I)
CC   is useful for diagnosing the onset or progression of BPH. (II) is useful
CC   for identifying an agent that modulates the onset or progression of BPH.
CC   The methods are useful to present information identifying the expression
CC   level in a tissue or cells, by comparing the expression level of genes
CC   given in the specification in the tissue or cells to the level of
CC   expression of gene in the database, and displaying the expression levels
CC   of at least one gene in the tissue or cell sample compared to the
CC   expression level in BPH. Agents using (II) are useful for treating BPH or
CC   prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CC   hyperplasia gene sequences of the invention
XX
SQ   Sequence 410 BP; 92 A; 96 C; 103 G; 119 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             |||||||||||||||  |||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:||||||||||||| ||||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             ||||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
             ||||||||||||||||||  |||||||||||||||||||||||:|||
Db          50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 230
ABN94040/c
ID   ABN94040 standard; DNA; 410 BP.
XX
AC   ABN94040;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #538 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 538; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 410 BP; 92 A; 96 C; 103 G; 119 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             |||||||||||||||  |||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:||||||||||||| ||||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             ||||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111
```

```
Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||||  |||||||||||||||||||||||||:|||
Db         50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 231
AAV43242
ID   AAV43242 standard; cDNA; 429 BP.
XX
AC   AAV43242;
XX
DT   29-DEC-1998  (first entry)
XX
DE   MSRV-1 clone 5 pol gene region from patient 46-7.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .429
FT                   /*tag=  a
FT                   /product= "Protein AAW71077"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71077.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 19; Page 203; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) pol gene region of clone 5 obtained from an MS
CC   patient 46-7. The example compares the pol regions of different clones
CC   obtained from the same patients plasma sample. The results show that the
CC   patient possesses a virion population which comprises of different MSRV
CC   variants. The invention provides complete or partial genomic sequences of
CC   the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC   these genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 429 BP; 131 A; 109 C; 86 G; 103 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 339;  DB 1;  Length 429;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  358;  Conservative   40;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| ||||||  ||||:|||:  |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATAGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| |||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |: :|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAGGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| | |||:|||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||||:||||||| |||||||||||||:|::|||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||||| | |:  | :||||||||||:||||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              | |||||||||||||||||||||||||||| ||:|||||||||:|:||||  :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 232
ADB84432
ID   ADB84432 standard; DNA; 429 BP.
XX
AC   ADB84432;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 pol region DNA from patient sample #3.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84436.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
```

```
XX
PS   Example 19; Fig 49A; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 429 BP; 131 A; 109 C; 86 G; 103 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 339;  DB 1;  Length 429;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  358;  Conservative   40;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:  |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATAGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
               |:||: ||::|||||||||||||||| |||||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |: :|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAGGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| | |||:|||||||||| ||||||:::|||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||| |||||||||||||:|::|||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||:|||||||||| | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              | |||||||||||||||||||||||||||| ||:|||||||||:|:||||  :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 233
AAI13454/c
ID   AAI13454 standard; DNA; 405 BP.
XX
AC   AAI13454;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3387 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 3387; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db         400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:|||||:|||||||||| ||||:|||:|||
Db         341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db         282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||  ||||||||||::|||||||||||||||||||||||
Db         222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db         162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
```

```
                |||||||||||||||| |||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||||:|||| |||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 234
ABA55151/c
ID   ABA55151 standard; DNA; 405 BP.
XX
AC   ABA55151;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3456.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 3456; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db         400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db         341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db         282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db         222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||||| |||||||||||||||||:|||||::|||||||||||||
Db         162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| |||||||||||:||||||||||||:||||||||||||||||||
Db         106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||||:|||| |||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 235
AAI34810/c
ID   AAI34810 standard; DNA; 405 BP.
XX
AC   AAI34810;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #3496 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 3496; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db         400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342
```

```
Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
Db          46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1
```

```
RESULT 236
ABA44696/c
ID   ABA44696 standard; DNA; 405 BP.
XX
AC   ABA44696;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3391.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 3391; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
Db          46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1
```

```
RESULT 237
ABA24898/c
ID   ABA24898 standard; DNA; 405 BP.
XX
AC   ABA24898;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #3364 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
```

```
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 3364; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||||| | ||||||||:||||||||||||:||||||||:|||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||||| ||||||||||||||||:|||||::||||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| |||||||||||:||||||||||:||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||||:|||| |||||||||||||||||||||||:|
Db          46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 238
AAK28864/c
ID   AAK28864 standard; DNA; 405 BP.
XX
AC   AAK28864;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3421.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 3421; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||||| | ||||||||:||||||||||||:||||||||:|||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||||| ||||||||||||||||:|||||::||||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| |||||||||||:||||||||||:||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||||:|||| |||||||||||||||||||||||:|
Db          46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 239
```

```
AAK03413/c
ID   AAK03413 standard; DNA; 405 BP.
XX
AC   AAK03413;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 3404.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 3404; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||| | ||||||||:|||||||||||:||||||||:|||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:|||||||| ||||||||||::||||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||| ||||||||||||||:|||||::||||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| |||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 240
ABS28480/c
ID   ABS28480 standard; DNA; 405 BP.
XX
AC   ABS28480;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 3470.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 3470; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
```

```
                ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
                |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
                ||:|||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
                :|:||||||||||:|||||||| ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
                |||||||    |||||||||| |||||||||||||||||:||||::|||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
                |||||||||||||||| |||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
                |:||||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 241
AAI03334/c
ID   AAI03334 standard; DNA; 405 BP.
XX
AC   AAI03334;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #3325 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 3325; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
                ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
                |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
                ||:|||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
                :|:||||||||||:|||||||| ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
                |||||||    |||||||||| |||||||||||||||||:||||::|||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
                |||||||||||||||| |||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
                |:||||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 242
ABS03396/c
ID   ABS03396 standard; DNA; 405 BP.
XX
AC   ABS03396;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 3387.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
```

```
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 3387; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 405 BP; 100 A; 93 C; 89 G; 123 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||||| | ||||||||:|||||||||||||:||||||||:|||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||||::|||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||||| ||||||||||||||||:|||||::||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:||||||||||||:|||| ||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 243
AAI14299/c
ID   AAI14299 standard; DNA; 409 BP.
XX
AC   AAI14299;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #4232 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 4232; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
```

```
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 244
ABA56026/c
ID   ABA56026 standard; DNA; 409 BP.
XX
AC   ABA56026;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #4331.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 4331; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 245
AAI35675/c
ID   AAI35675 standard; DNA; 409 BP.
XX
AC   AAI35675;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #4361 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
```

```
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 4361; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
               |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 246
ABA45530/c
ID   ABA45530 standard; DNA; 409 BP.
XX
AC   ABA45530;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #4225.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 4225; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
               |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              ||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 247
ABA25687/c
```

```
ID   ABA25687 standard; DNA; 409 BP.
XX
AC   ABA25687;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #4153 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 4153; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| ||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| |||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 248
AAK29717/c
ID   AAK29717 standard; DNA; 409 BP.
XX
AC   AAK29717;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 4274.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 4274; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| ||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290
```

```
Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              ||||||| |||||||| |||||| ||||||||||||||||  | |||||  ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 249
AAK04230/c
ID   AAK04230 standard; DNA; 409 BP.
XX
AC   AAK04230;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 4221.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 4221; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||| ||||| |||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              ||||||| |||||||| |||||| ||||||||||||||||  | |||||  ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 250
ABS29354/c
ID   ABS29354 standard; DNA; 409 BP.
XX
AC   ABS29354;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 4344.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 4344; 658pp; English.
```

```
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              ||||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 251
AAI04133/c
ID   AAI04133 standard; DNA; 409 BP.
XX
AC   AAI04133;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #4124 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 4124; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              ||||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 252
ABS04271/c
ID   ABS04271 standard; DNA; 409 BP.
XX
AC   ABS04271;
XX
DT   19-AUG-2002  (first entry)
```

```
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 4262.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 4262; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 409 BP; 103 A; 87 C; 134 G; 85 T; 0 U; 0 Other;

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| ||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              ||||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | |||||  |||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 253
AAI14380/c
ID   AAI14380 standard; DNA; 414 BP.
XX
AC   AAI14380;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #4313 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
```

```
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 4313; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | |||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| ||||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 254
ABA56109/c
ID   ABA56109 standard; DNA; 414 BP.
XX
AC   ABA56109;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #4414.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 4414; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | |||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| ||||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 255
AAI35756/c
ID   AAI35756 standard; DNA; 414 BP.
XX
AC   AAI35756;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #4442 used to measure gene expression in human placenta sample.
```

```
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 4442; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
              |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
              ||||||| |||||||||||||||| | |||||| |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
              ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
              |||||||| ||| |||| |||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db          55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15

RESULT 256
ABA45604/c
ID   ABA45604 standard; DNA; 414 BP.
XX
AC   ABA45604;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #4299.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 4299; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
              |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295
```

```
Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  ||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15
```

RESULT 257
ABA25761/c
ID ABA25761 standard; DNA; 414 BP.
XX
AC ABA25761;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4227 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 1; SEQ ID NO 4227; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

```
  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| ||||||||||||||||| | ||||||  ||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15
```

RESULT 258
AAK29793/c
ID AAK29793 standard; DNA; 414 BP.
XX
AC AAK29793;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 4350.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.

```
XX
PS   Example 4; SEQ ID NO 4350; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 259
AAK04297/c
ID   AAK04297 standard; DNA; 414 BP.
XX
AC   AAK04297;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 4288.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 4288; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 260
ABS29438/c
ID   ABS29438 standard; DNA; 414 BP.
XX
AC   ABS29438;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 4428.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
```

```
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 4428; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | |||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             |||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15

RESULT 261
AAI04204/c
ID   AAI04204 standard; DNA; 414 BP.
XX
AC   AAI04204;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #4195 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 4195; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | |||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
```

```
                ||||||||:||| |||| ||||||||||||  |:||:||||:|  |:||||  :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
                |||||||  ||  ||||||||  ||||||  |||||||||||||||:|||:||||  ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
                ||||||||  |||  ||||  ||||||||||||||||||||||||||  ||  |||:|||  ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
                ||||||  |||||  ||||||||||  ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 262
ABS04353/c
ID   ABS04353 standard; DNA; 414 BP.
XX
AC   ABS04353;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 4344.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 4344; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 414 BP; 111 A; 74 C; 122 G; 107 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
                ||| | ||||||||||||||||||||||||||||||||||||||||  |||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
                ||||||  |||||  |||  ||||||||||||||||||||  ||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
                |||||||  ||||||||||||||||  | ||||||  ||||||||||||||  |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
                ||||||||:||| |||| ||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
                |||||||  ||  ||||||||  ||||||  |||||||||||||||:|||:||||  ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
                ||||||||  |||  ||||  ||||||||||||||||||||||||||  ||  |||:|||  ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
                ||||||  |||||  ||||||||||  ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 263
AAV09123
ID   AAV09123 standard; cDNA; 374 BP.
XX
AC   AAV09123;
XX
DT   28-MAY-1998  (first entry)
XX
DE   5' nucleotide sequence of clone AJ172_2.
```

```
XX
KW   AJ172_2 clone; cDNA library; human adult testes; cell proliferation;
KW   cell differentiation; ss.
XX
OS   Homo sapiens.
XX
PN   US5708157-A.
XX
PD   13-JAN-1998.
XX
PF   26-JUL-1996;   96US-00686878.
XX
PR   26-JUL-1996;   96US-00686878.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Mccoy JM,  Treacy M,  Racie LA,  Merberg D,  Spaulding V,  Jacobs K;
PI   Lavallie ER,  Evans C;
XX
DR   WPI; 1998-109352/10.
XX
PT   Human testes derived cDNA and protein(s) - may be useful for laboratory
PT   use, or for stimulating cell proliferation and differentiation.
XX
PS   Disclosure; Col 79-80; 48pp; English.
XX
CC   This sequence is the 5' nucleotide sequence for the AJ172-2 clone, which
CC   is isolated from a human adult testes cDNA library. When the AJ172_2
CC   sequence was searched in the Genbank database, the clone demonstrated at
CC   least some homology with the following sequences: Friend murine leukemiz
CC   virus; Moloney murine leukemia virus genome; pol protein, Gibbon leukemia
CC   virus; and an EST identified as "yh46a09.s1 Homo sapiens cDNA clone
CC   132760 3'"
XX
SQ   Sequence 374 BP; 92 A; 100 C; 80 G; 80 T; 0 U; 22 Other;

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              |  ||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              |  |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 264
AAV15135
ID   AAV15135 standard; cDNA; 374 BP.
XX
AC   AAV15135;
XX
DT   29-JUN-1998  (first entry)
XX
DE   Human adult testes cDNA clone AJ172_2 5' cDNA portion.
XX
KW   Human; adult; testes; AJ172_2; secreted protein; cytokine; nutrition;
KW   cell proliferation; differentiation; tumour; immune; ds.
XX
OS   Homo sapiens.
XX
PN   WO9804696-A1.
XX
PD   05-FEB-1998.
XX
PF   23-JUL-1997;   97WO-US012910.
XX
PR   26-JUL-1996;   96US-00686878.
PR   27-SEP-1996;   96US-00721489.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Racie LA,  Merberg D,  Treacy M;
PI   Spaulding V;
XX
DR   WPI; 1998-130687/12.
XX
PT   Polynucleotides and secreted proteins having cytokine and cell
PT   proliferation/differentiation activity - isolated from a human adult
PT   testes cDNA library.
XX
PS   Claim 13; Page 33; 48pp; English.
XX
CC   The present sequence represents the 5' cDNA portion of the clone AJ172_2
CC   from a human adult testes cDNA library. The nucleotide sequences and
CC   protein from clone AJ172_2, have uses in nutrition, cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis regulating activity, chemotactic/chemokinetic
CC   activity, haemostatic or thrombotic activity, receptor/ligand activity,
CC   anti-inflammatory activity, tumour inhibition activity or other
CC   activities. The products can also be used for production of antibodies
CC   and for detection and diagnosis
XX
SQ   Sequence 374 BP; 92 A; 100 C; 80 G; 80 T; 0 U; 22 Other;

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              |  ||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              |  |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 265
AEA20777
ID   AEA20777 standard; cDNA; 415 BP.
```

XX
AC AEA20777;
XX
DT 11-AUG-2005 (first entry)
XX
DE Novel human polynucleotide SEQ ID NO 1471.
XX
KW vulnerary; CNS-gen.; gene therapy; diagnostic; forensic; mapping;
KW DNA purification; protein purification; osteoarthritis; antiarthritic;
KW osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;
KW periodontal disease; antiinflammatory; mouth disease; burns; injury;
KW peripheral neuropathy; Alzheimers disease; neuroprotective; nootropic;
KW degeneration; parkinsons disease; antiparkinsonian; neurological disease;
KW cerebrovascular ischemia; cerebroprotective; vasotropic;
KW cardiovascular disease; autoimmune disease; immunosuppressive;
KW immune disorder; viral infection; virucide; infection; cancer;
KW cytostatic; neoplasm; ss.
XX
OS Homo sapiens.
XX
PN WO2005049806-A2.
XX
PD 02-JUN-2005.
XX
PF 11-MAR-2004; 2004WO-US007412.
XX
PR 14-MAR-2003; 2003US-00389559.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;
PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;
PI Wehrman T, Weng G, Boyle B;
XX
DR WPI; 2005-417730/42.
XX
PT New polynucleotide encoding a polypeptide with biological activity,
PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,
PT CNS and peripheral disease, stroke, autoimmune disorders, viral
PT infection, or cancer.
XX
PS Example 3; SEQ ID NO 1471; 500pp; English.
XX
CC The invention describes a new isolated polynucleotide (I) encoding a
CC polypeptide with biological activity comprising: a nucleotide sequence of
CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes
CC to the sequence of (i) under stringent hybridization conditions; or a
CC nucleotide sequence having greater than 99% sequence identity with the
CC sequence of (i). Also described are: a(n) (expression)vector comprising
CC (I); a host cell genetically engineered to comprise (I) operatively,
CC associated with a regulatory sequence that modulates expression of the
CC polynucleotide in the host cell; an isolated polypeptide comprising a
CC sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide
CC is: a polypeptide encoded by (I); or a polypeptide encoded by a
CC polynucleotide hybridizing under stringent conditions with any one of SEQ
CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a
CC carrier; an antibody directed against the polypeptide of (3); a method
CC for detecting (I) in a sample; a method for detecting the polypeptide of
CC (3) in a sample; a method for identifying a compound that binds to the
CC polypeptide of (3); a method of producing the polypeptide of (3); and a
CC collection of polynucleotides, where the collection comprising of at
CC least one of SEQ ID NOS: 1-567. (I) is a polynucleotide comprising any of
CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological
CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:
CC 568-1134. All sequences are fully defined in the specification. The
CC sequences and methods are useful in diagnostics, forensic, and gene
CC mapping, in identifying of mutations responsible for genetic disorders or
CC other traits, in assessing biodiversity, and for producing many other
CC types of data and products dependent on DNA and amino acid sequences. The
CC composition and method are useful for treating a disease or disorder,
CC e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and
CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,
CC autoimmune disorders, viral infection, or cancer. This sequence
CC represents a novel polynucleotide of the invention.
XX
SQ Sequence 415 BP; 83 A; 137 C; 102 G; 93 T; 0 U; 0 Other;

Query Match 4.0%; Score 300.7; DB 1; Length 415;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 352; Conservative 4; Mismatches 40; Indels 23; Gaps 2;

```
Qy      151 GCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGA 210
            |||||||||||||||| ||||||||||||||||||| | |    |||| ||| || |||||
Db        1 GCTGAGCTTTTGCTCGCCGTCCACCACTGCTGTTTGTCGCTGTCGCAGACCCGCTGCTGA 60

Qy      211 CTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCA 270
            || |||||| |||||||||| |||||||||| |||||||||||||||||  ||:|| ||||
Db       61 CTGCCATCCGTCTGGATCCGGCAGGGTGTCCACTGTGCTCCTGATCCAATGAGGCACCCA 120

Qy      271 TTGCCGCTCCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGT 330
            |||||||||||  ||  |||||||||||||||||||| ||||||| |||| ||||| ||||
Db      121 TTGCCGCTCCCGATCCGGCTAAAGGCTTGCCATTGTTCCTGCACAGCTAGGTGCCCGGGT 180

Qy      331 TTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACG 390
            |||| |||||  |||||||| ||| |||||||||||| |||||||||  |||| |||||
Db      181 TTGTCCTAATCCAGCTGAACGCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGATCCACG 240

Qy      391 GCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGT 450
            ||||||||||:| | |||||||| |||||||||||||||||||||||||||||||||| |
Db      241 GCTTCTAATAGAGCGATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCAT 300

Qy      451 GAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTG 510
            ||||:||| |||| |||||||||||||| || |:|||||||||||||||||||||||||
Db      301 GAGGCCAA-GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCATCTTGGAAGCGGCC-- 357

Qy      511 CTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCG 569
                                  | |||||||||||||||||| ||||||||||||||||
Db      358 --------------------CGCCACCATCTTGGGAGCTCTGGGAGCAAGGACCCCCCG 396
```

RESULT 266
AEA20798
ID AEA20798 standard; cDNA; 415 BP.
XX
AC AEA20798;
XX
DT 11-AUG-2005 (first entry)
XX
DE Novel human polynucleotide SEQ ID NO 1492.
XX
KW vulnerary; CNS-gen.; gene therapy; diagnostic; forensic; mapping;
KW DNA purification; protein purification; osteoarthritis; antiarthritic;
KW osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;
KW periodontal disease; antiinflammatory; mouth disease; burns; injury;
KW peripheral neuropathy; Alzheimers disease; neuroprotective; nootropic;
KW degeneration; parkinsons disease; antiparkinsonian; neurological disease;
KW cerebrovascular ischemia; cerebroprotective; vasotropic;
KW cardiovascular disease; autoimmune disease; immunosuppressive;
KW immune disorder; viral infection; virucide; infection; cancer;
KW cytostatic; neoplasm; ss.
XX
OS Homo sapiens.
XX
PN WO2005049806-A2.
XX
PD 02-JUN-2005.
XX
PF 11-MAR-2004; 2004WO-US007412.
XX
PR 14-MAR-2003; 2003US-00389559.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;
PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;
PI Wehrman T, Weng G, Boyle B;

```
XX
DR   WPI; 2005-417730/42.
XX
PT   New polynucleotide encoding a polypeptide with biological activity,
PT   useful for treating a disease or disorder, e.g. osteoarthritis, burns,
PT   CNS and peripheral disease, stroke, autoimmune disorders, viral
PT   infection, or cancer.
XX
PS   Example 3; SEQ ID NO 1492; 500pp; English.
XX
CC   The invention describes a new isolated polynucleotide (I) encoding a
CC   polypeptide with biological activity comprising: a nucleotide sequence of
CC   SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes
CC   to the sequence of (i) under stringent hybridization conditions; or a
CC   nucleotide sequence having greater than 99% sequence identity with the
CC   sequence of (i). Also described are: a(n) (expression)vector comprising
CC   (I); a host cell genetically engineered to comprise (I) operatively,
CC   associated with a regulatory sequence that modulates expression of the
CC   polynucleotide in the host cell; an isolated polypeptide comprising a
CC   sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide
CC   is: a polypeptide encoded by (I); or a polypeptide encoded by a
CC   polynucleotide hybridizing under stringent conditions with any one of SEQ
CC   ID NOS: 1-567; a composition comprising the polypeptide of (3) and a
CC   carrier; an antibody directed against the polypeptide of (3); a method
CC   for detecting (I) in a sample; a method for detecting the polypeptide of
CC   (3) in a sample; a method for identifying a compound that binds to the
CC   polypeptide of (3); a method of producing the polypeptide of (3); and a
CC   collection of polynucleotides, where the collection comprising of at
CC   least one of SEQ ID NOS: 1-567. (I) is a polynucleotide comprising any of
CC   the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological
CC   activity, which comprises any of the amino acid sequence of SEQ ID NOS:
CC   568-1134. All sequences are fully defined in the specification. The
CC   sequences and methods are useful in diagnostics, forensic, and gene
CC   mapping, in identifying of mutations responsible for genetic disorders or
CC   other traits, in assessing biodiversity, and for producing many other
CC   types of data and products dependent on DNA and amino acid sequences. The
CC   composition and method are useful for treating a disease or disorder,
CC   e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and
CC   peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,
CC   autoimmune disorders, viral infection, or cancer. This sequence
CC   represents a novel polynucleotide of the invention.
XX
SQ   Sequence 415 BP; 83 A; 137 C; 102 G; 93 T; 0 U; 0 Other;

  Query Match             4.0%;  Score 300.7;  DB 1;  Length 415;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  352;  Conservative    4;  Mismatches   40;  Indels   23;  Gaps    2;

Qy        151 GCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGA 210
              |||||||||||||||| |||||||||||||||||||| | |    |||| ||| || |||||
Db          1 GCTGAGCTTTTGCTCGCCGTCCACCACTGCTGTTTGTCGCTGTCGCAGACCCGCTGCTGA 60

Qy        211 CTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCA 270
              || |||||| ||||||||| |||||||||| |||||||||||||||  ||:|| ||||
Db         61 CTGCCATCCGTCTGGATCCGGCAGGGTGTCCACTGTGCTCCTGATCCAATGAGGCACCCA 120

Qy        271 TTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGT 330
              |||||||||| ||  |||||||||||||||||||| ||||||| |||| ||||| ||||
Db        121 TTGCCGCTCCCGATCCGGCTAAAGGCTTGCCATTGTTCCTGCACAGCTAGGTGCCCGGGT 180

Qy        331 TTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACG 390
              |||| |||||  |||||||| ||| |||||||||||| |||||||||  |||| |||||
Db        181 TTGTCCTAATCCAGCTGAACGCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGATCCACG 240

Qy        391 GCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGT 450
              ||||||||||:| | |||||||| |||||||||||||||||||||||||||||||| |
Db        241 GCTTCTAATAGAGCGATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCAT 300

Qy        451 GAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTG 510
              ||||:||| |||| |||||||||||||| || |:|||||||||||||||||||||||||
Db        301 GAGGCCAA-GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCATCTTGGAAGCGGCC-- 357

Qy        511 CTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCG 569
```

```
                                                  | |||||||||||||||||| ||||||||||||||||
Db        358 --------------------CGCCACCATCTTGGGAGCTCTGGGAGCAAGGACCCCCCG 396
```

```
RESULT 267
ABA73835/c
ID   ABA73835 standard; DNA; 336 BP.
XX
AC   ABA73835;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #22140.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 22140; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:||||||||||  |||||||||||||||:|||||::|||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||: |||||||||||:|||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:||||  ||| |||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97
```

```
Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db         96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
              |||||||||||||||||||| |: | ||||||||||
Db         37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 268
AAI54279/c
ID   AAI54279 standard; DNA; 336 BP.
XX
AC   AAI54279;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #22965 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 22965; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:|||||||||||  |||||||||||||||||:||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: |||||||||||:|||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:||||  ||| |||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db         96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
              |||||||||||||||||||| |: | ||||||||||
Db         37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 269
ABA38992/c
ID   ABA38992 standard; DNA; 336 BP.
XX
AC   ABA38992;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #17458 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 17458; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;
```

```
Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:||||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||||: ||||||||||:||||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             |||||||  |:||: ||||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1
```

RESULT 270
AAK48452/c
ID AAK48452 standard; DNA; 336 BP.
XX
AC AAK48452;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 23009.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 23009; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

Query Match 3.9%; Score 292.1; DB 1; Length 336;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 302; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:||||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||||: ||||||||||:||||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             |||||||  |:||: ||||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1
```

RESULT 271
AAK22286/c
ID AAK22286 standard; DNA; 336 BP.
XX
AC AAK22286;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 22277.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 22277; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human

```
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 272
ABS48146/c
ID   ABS48146 standard; DNA; 336 BP.
XX
AC   ABS48146;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 23136.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 23136; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 273
ABS22263/c
ID   ABS22263 standard; DNA; 336 BP.
XX
AC   ABS22263;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 22254.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
```

PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 22254; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 336 BP; 86 A; 72 C; 77 G; 101 T; 0 U; 0 Other;

Query Match 3.9%; Score 292.1; DB 1; Length 336;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 302; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||||: |||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| |||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1
```

RESULT 274
AAI27830/c
ID AAI27830 standard; DNA; 387 BP.
XX
AC AAI27830;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #17763 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 17763; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

```
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| |||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    |||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 275
ABA76145/c
ID   ABA76145 standard; DNA; 387 BP.
XX
AC   ABA76145;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #24450.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 24450; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| |||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    |||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 276
AAI56806/c
ID   AAI56806 standard; DNA; 387 BP.
XX
AC   AAI56806;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #25492 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

```
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 25492; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| ||||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              ||||  |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 277
AAK24809/c
ID   AAK24809 standard; DNA; 387 BP.
XX
AC   AAK24809;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 24800.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 24800; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| ||||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              ||||  |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 278
ABS50402/c
ID   ABS50402 standard; DNA; 387 BP.
XX
AC   ABS50402;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 25392.
```

```
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 25392; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||| |||||| ||||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| |||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 279
ABS24297/c
ID   ABS24297 standard; DNA; 387 BP.
XX
AC   ABS24297;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 24288.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 24288; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
```

```
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 387 BP; 91 A; 90 C; 131 G; 75 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| |||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    |||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 280
AAA98187
ID   AAA98187 standard; DNA; 341 BP.
XX
AC   AAA98187;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-T47D-W2.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 58; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 341 BP; 98 A; 82 C; 84 G; 77 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 283.3;  DB 1;  Length 341;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  314;  Conservative    8;  Mismatches   10;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||||||||||||||||||| |||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTCGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  |||||||         |||| |||||||||||||||||||||||||||||||
Db         121 CACACCCAACCAATCAGGTAGTAAAGAGGGCTCACTAAAATGCTAATTAGGCAAAGACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||||:|||||:|||||||||||||||||||||||||||||||||
Db         181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||||||:||||||||||||  |||  |||  ||||||||||||||||||||
Db         241 GGATATAAACCCAAGTCTTCGAGCCGGCAATGGCTACCTTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||          |||||||||||||||||||||:||||
Db         301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 341
```

```
RESULT 281
AAA98200
ID   AAA98200 standard; DNA; 343 BP.
XX
AC   AAA98200;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W8.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 61; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 105 A; 84 C; 81 G; 73 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 270.3;  DB 1;  Length 343;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  307;  Conservative    7;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||| |||||||||| |||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAAACAGGACTAGCAGGATTTCCTAGGCCGATTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :||||||||| ||||||| |||||||||||||||||||||||||||||||||||| ||:|
Db         61 CTAAGCCTAGATGGGAAGTTGACCACATCCACCTTTAAACACGGGGCTTGCAACTCAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              |||||| |||| |||         ||||||| | |||||||||||||||||||| ||||
Db         121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||||||:||||||:|||||||||||| ||||||||||||||||
Db         181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  |||||||||||||| | |||| |||||||||||||||||||||
Db         241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||          |||||||||||||||||||||:||||:
Db         301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 282
AAA98198
ID   AAA98198 standard; DNA; 343 BP.
XX
AC   AAA98198;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W5.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 60; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
```

```
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 104 A; 84 C; 80 G; 75 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 269.9;  DB 1;  Length 343;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  306;  Conservative    8;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||  |:||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCAACTAAGAATCT 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||| ||||||||||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| |||| |||          ||||||| | |||||||||||||||||||| ||||
Db            121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||| |||||||:||||||:|||||||||||| ||||||||||||||||
Db           181 GAGGTAAAGAAATAGTCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||||| |  |||||||||||||||| | |||| ||||||||||||||||||||
Db           241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||          |||||||||||||||||||||:||||:
Db           301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 283
AAI13385/c
ID   AAI13385 standard; DNA; 379 BP.
XX
AC   AAI13385;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3318 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 3318; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             |||||||||||  ||   ||| |      ||  ||||||||||||||||||| | |
Db            379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| ||||||||| ||||||||||
Db            323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db            263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| |||||||||| ||||||||||||
Db            204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |  | |||||||||||||||||||||| ||||  |||||||| |
Db            145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
             |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db             85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
             | | | |||||||| || | |
Db            25 GGACTAGCCCCTGCAATACTCC 4


RESULT 284
ABA55083/c
ID   ABA55083 standard; DNA; 379 BP.
XX
AC   ABA55083;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3388.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
```

```
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 3388; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             ||||||||||||  ||||  |||||||||||||||||||||  |||||||||  ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||  |||||||  | | |||||  ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  ||||||||||||||||||||  |||  ||||||||||  ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |   | ||||||||||||||||||||||||| ||||  ||||||||
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              ||||||||||  |||||||| |||||  |||||  |||||||||||||||  || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  |||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 285
AAI34739/c
ID   AAI34739 standard; DNA; 379 BP.
XX
AC   AAI34739;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #3425 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 3425; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             ||||||||||||  |||| ||||||||||||||||||||  |||||||||  ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||  |||||||  | | |||||  ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  ||||||||||||||||||||  |||  ||||||||||  ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |   | ||||||||||||||||||||||||| ||||  ||||||||
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              ||||||||||  |||||||| |||||  |||||  |||||||||||||||  || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  |||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 286
ABA44631/c
ID   ABA44631 standard; DNA; 379 BP.
XX
```

```
AC   ABA44631;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3326.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 3326; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             ||||||||||||| |||| |||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  ||||||||||||||||||||| ||| |||||||||| ||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |  | |||||||||||||||||||||||| ||||  |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 287
AAK28798/c
ID   AAK28798 standard; DNA; 379 BP.
XX
AC   AAK28798;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3355.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 3355; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
```

```
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
              |||||||||||| |||| ||||||||||||||||||||| ||||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
              |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
              |||||||||  |||||||||||||||||||| ||| |||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
              |||||||||||||  |   | |||||||||||||||||||||||| ||||  |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              ||||||||||  |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4
```

RESULT 288
AAK03348/c
ID AAK03348 standard; DNA; 379 BP.
XX
AC AAK03348;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 3339.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 3339; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

```
  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
              ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
              |||||||||||| |||| ||||||||||||||||||||| ||||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
              |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
              |||||||||  |||||||||||||||||||| ||| |||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
              |||||||||||||  |   | |||||||||||||||||||||||| ||||  |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              ||||||||||  |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4
```

RESULT 289
ABS28403/c
ID ABS28403 standard; DNA; 379 BP.
XX
AC ABS28403;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 3393.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.

```
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 3393; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||    ||| |       ||  ||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| ||||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| ||||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             ||||||||||||||  |  |||||||||||||||||||||||| ||||  |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
             |||||||||| ||||||||  ||||| ||||| ||||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
             | | |  |||||||  || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 290
AAI03278/c
ID   AAI03278 standard; DNA; 379 BP.
XX
AC   AAI03278;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #3269 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 3269; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||    ||| |       ||  ||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| ||||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| ||||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             ||||||||||||||  |  |||||||||||||||||||||||| ||||  |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
             |||||||||| ||||||||  ||||| ||||| ||||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
             | | |  |||||||  || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 291
ABS03326/c
```

```
ID   ABS03326 standard; DNA; 379 BP.
XX
AC   ABS03326;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 3317.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 3317; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 379 BP; 87 A; 77 C; 92 G; 123 T; 0 U; 0 Other;

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
              ||||||||||||  ||   ||| |       ||  |||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
              ||||||||||||  ||||  |||||||||||||||||||||  |||||||||  |||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
              |||||||||||||||||||||||  |||||||  | |  |||||  ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
              |||||||||  ||||||||||||||||||| ||| ||||||||||  ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
              |||||||||||||  |  | ||||||||||||||||||||||||  ||||  ||||||||  |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
                |||||||||  ||||||||  |||||  |||||  ||||||||||||||  ||  |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  |||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 292
AAA98194
ID   AAA98194 standard; DNA; 343 BP.
XX
AC   AAA98194;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W2.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
```

```
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 59; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 104 A; 84 C; 77 G; 78 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 265.1;  DB 1;  Length 343;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  303;  Conservative    8;  Mismatches   22;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         ||||||||  |||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||| ||||||||||||||:||| ||:|||||||||||||||||||||||||||||||
Db        181 GAGGTAGAGAAATAGCCAATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCC 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||||||  |||||||| |||| ||| |||| ||||| ||||||||||||||
Db        241 GGATATAAACCCAAGCATTCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||| |||||          |||||||||||||||||:||||:
Db        301 TATGGGAGCTCTATTTTC----------ACTCTATTAAATCTTGCAACTGCA 342


RESULT 293
AAA98188
ID   AAA98188 standard; DNA; 341 BP.
XX
AC   AAA98188;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-T47D-W4.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 58; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 341 BP; 98 A; 80 C; 82 G; 81 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.9;  DB 1;  Length 341;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  302;  Conservative    7;  Mismatches   23;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||:||||||| |
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATTC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| |||||| |||||||| |||||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             || | ||||||||||         |||||||||||||||||||||||||||| || ||||
Db        121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||| |||||| ||||||||||:| ||||:||| |||||||||||||||||||||||||||
Db        181 GAGGCAAAGAAGTAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGACAATGATCG 240
```

```
Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             ||||||||||||| |  |||||||| || || ||  |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||          ||||||||||||||||||||:||||
Db        301 TATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGC 341


RESULT 294
AAI21066
ID   AAI21066 standard; DNA; 310 BP.
XX
AC   AAI21066;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #10999 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 10999; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             |||||  || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||||||| || | |||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||||| | |||||||||||||||||||||||| |||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||
Db        300 AGCAGGACTGA 310


RESULT 295
ABA66142
ID   ABA66142 standard; DNA; 310 BP.
XX
AC   ABA66142;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #14447.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 14447; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
```

```
                |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
                ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
                |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                |||||||||||||||||||||||| | ||||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
                ||||||||||
Db        300 AGCAGGACTGA 310
```

RESULT 296
AAI46325
ID AAI46325 standard; DNA; 310 BP.
XX
AC AAI46325;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #15011 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 15011; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

Query Match 3.5%; Score 262.7; DB 1; Length 310;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 287; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

```
Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
                |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
                ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
                |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                |||||||||||||||||||||||| | ||||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
                ||||||||||
Db        300 AGCAGGACTGA 310
```

RESULT 297
ABA48260
ID ABA48260 standard; DNA; 310 BP.
XX
AC ABA48260;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #6955.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 4; SEQ ID NO 6955; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived

```
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             ||||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 298
ABA33210
ID   ABA33210 standard; DNA; 310 BP.
XX
AC   ABA33210;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #11676 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 11676; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             ||||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 299
AAK40307
ID   AAK40307 standard; DNA; 310 BP.
XX
AC   AAK40307;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 14864.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
```

```
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 14864; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||||| || | ||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 300
AAK14563
ID   AAK14563 standard; DNA; 310 BP.
XX
AC   AAK14563;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 14554.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 14554; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||||| || | ||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 301
ABS39877
ID   ABS39877 standard; DNA; 310 BP.
XX
```

```
AC   ABS39877;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 14867.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 14867; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
              |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
              | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              ||||||||||
Db        300 AGCAGGACTGA 310


RESULT 302
AAI06786
ID   AAI06786 standard; DNA; 310 BP.
XX
AC   AAI06786;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #6777 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 6777; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59
```

```
Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
              | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||| | ||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310
```

RESULT 303
ABS14331
ID ABS14331 standard; DNA; 310 BP.
XX
AC ABS14331;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 14322.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 14322; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310 BP; 98 A; 73 C; 85 G; 54 T; 0 U; 0 Other;

Query Match 3.5%; Score 262.7; DB 1; Length 310;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 287; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

```
Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
              | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             |||||||||||||||||||||||| | ||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
             ||||||||||||
Db        300 AGCAGGACTGA 310
```

RESULT 304
AAA98192
ID AAA98192 standard; DNA; 343 BP.
XX
AC AAA98192;
XX
DT 30-JAN-2001 (first entry)
XX

```
DE   Human retrovirus LTR DNA fragment HERV-W11.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 59; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 101 A; 80 C; 84 G; 78 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| |||||||||||||||| :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:||||||  ||||||| |||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         || |||||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||| |||||||||:||| ||:|||||||| |||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||| ||||          ||||||||||||||||||||:||||:
Db        301 TATGGGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 305
AAA98197
ID   AAA98197 standard; DNA; 343 BP.
XX
AC   AAA98197;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W4.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 60; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 102 A; 79 C; 83 G; 79 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
```

```
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              ||||||||||||||||||||||||||||||||| ||||||||||||||| :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| ||||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              |||||| ||||||||         || |||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||| |||||||||:||| ||:|||||||||| ||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              ||||| ||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 306
AAA98193
ID   AAA98193 standard; DNA; 343 BP.
XX
AC   AAA98193;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W18.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 59; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 102 A; 80 C; 83 G; 78 T; 0 U; 0 Other;

  Query Match             3.5%;  Score 258.7;  DB 1;  Length 343;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  299;  Conservative    8;  Mismatches   26;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              ||||||||||||||||||||||||||||||||| ||||||||||||||| | :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCCGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| ||||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              |||||| ||||||||         || |||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||| |||||||||:||| ||:|||||||||| ||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              ||||| ||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 307
AAA98195
ID   AAA98195 standard; DNA; 343 BP.
XX
AC   AAA98195;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W22.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
```

```
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 59; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 104 A; 76 C; 81 G; 82 T; 0 U; 0 Other;

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||| |||||| |||||||| |||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             || | ||||||||||         ||||||||  |||||||||||||||||||| ||||
Db        121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||         ||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 308
AAA98196
ID   AAA98196 standard; DNA; 343 BP.
XX
AC   AAA98196;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W23.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 60; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 104 A; 82 C; 79 G; 78 T; 0 U; 0 Other;

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         ||||||||  |||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240
```

```
Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 309
AAA98190
ID   AAA98190 standard; DNA; 343 BP.
XX
AC   AAA98190;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W1.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 58; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 106 A; 81 C; 79 G; 77 T; 0 U; 0 Other;

  Query Match             3.4%;  Score 253.9;  DB 1;  Length 343;
  Best Local Similarity   84.1%;  Pred. No. 0;

  Matches  296;  Conservative    8;  Mismatches   29;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| ||| |||||||||||||| ||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAAACATGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  |||||||          ||||||||  |||||||||||||||||||||| ||||
Db        121 CACACCCAACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||  ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||          |||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 310
AAA98199
ID   AAA98199 standard; DNA; 343 BP.
XX
AC   AAA98199;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W6.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 60; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
```

```
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 105 A; 83 C; 78 G; 77 T; 0 U; 0 Other;

  Query Match             3.4%;  Score 252.3;  DB 1;  Length 343;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  295;  Conservative    8;  Mismatches   30;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         ||||||||  ||||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 311
AAS84198
ID   AAS84198 standard; cDNA; 323 BP.
XX
AC   AAS84198;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20002.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG20011.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 20002; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 323 BP; 122 A; 63 C; 82 G; 56 T; 0 U; 0 Other;

  Query Match             3.4%;  Score 251.2;  DB 1;  Length 323;
  Best Local Similarity   77.4%;  Pred. No. 0;
  Matches  278;  Conservative   23;  Mismatches   22;  Indels   36;  Gaps    1;

Qy       4008 GAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAG 4067
             ||||||||||:||||:|||||||||||||||:|||:| |||||| ||||||||||||||
Db          1 GAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACCAGAATTAGGAGAAG 60

Qy       4068 RAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATG 4127
             :||||||||:|||||||||| || |||
Db         61 GAAAAAGGGTAAATATATATGCAAACT---------------------------------- 87

Qy       4128 MRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGG 4187
                 ||||| |:|:|||||||||||||||||||: ||:|||||| | |||:||||||||
Db         88 ---CAATAGGGAGAGAAAGGGAATTCCTAACTTCCAAGGGAACACTTGTCAAACATCAGG 144

Qy       4188 AAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACT 4247
             |||||||| ||:|||||||:||||:||||||||||||:||||||||:||||||||||||
Db        145 AAGCCATTAAGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACT 204

Qy       4248 GCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTG 4307
             ||:|||||||||| |||||||| :||||||||||:|||:|||: ||||||| :||||||
Db        205 GCTGGGGTCATCAGAAAGGAAAAGAAAGGGAAATAGAAGAGAACCGCCAAGCGGATATTG 264

Qy       4308 AAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
             ||||:| |||||| ||||||||||||||||||||||||||||||||    ||    ||| ||
Db        265 AAGCCAGAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGCTTACAGAAGGACCCCTA 323


RESULT 312
AAA98191
ID   AAA98191 standard; DNA; 343 BP.
XX
AC   AAA98191;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-W10.
XX
```

```
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 59; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 343 BP; 105 A; 83 C; 81 G; 74 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 250.7;  DB 1;  Length 343;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  294;  Conservative    8;  Mismatches   31;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||| ||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTCCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             || |||  || ||||         |||||||||||||||| || ||||||| || ||||
Db        121 CATACCCAACAAATCAGGTAGTAAAGAGAGCTCACTAAAATACTGATTAGGCGAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             ||||||| |||| ||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAGGAAACAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| |||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 313
ACH83988
ID   ACH83988 standard; DNA; 333 BP.
XX
AC   ACH83988;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #17183.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 17183; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
```

```
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 333 BP; 111 A; 84 C; 74 G; 64 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 247.3;  DB 1;  Length 333;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  301;  Conservative    0;  Mismatches   32;  Indels    9;  Gaps    3;

Qy       2328 GAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACC 2387
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACC 60

Qy       2388 CGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGAC 2447
               ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db         61 TGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCCGAAGTCTGGGCAACAGAAGGAC 119

Qy       2448 AATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCC 2507
               |||| ||||||| |||||||||| || |||||||||||||||||||||||| | |||
Db        120 AATACGGACGAG-CAAAGAATGCCTGTGCTGTTCAAGTTAAACTAAAGGATTCCGCCTCC 178

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
              |||||||||||||||||||||||||| |||||  |||||||||||||| ||||||||||
Db        179 TTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACTCCAAAAGATT 238

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
              ||||||||| |||||||||| ||| |||||||||||||||  |  ||||||| || | ||
Db        239 GTTAAGGACCTAAAAGCCCATGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC- 297

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAG 2669
                  |||   |  | ||||||||||   ||||||||||
Db        298 ------AATTTTAGGAGTACAGAAACCCAACAGACAGTGGAG 333


RESULT 314
AAI23208
ID   AAI23208 standard; DNA; 275 BP.
XX
AC   AAI23208;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13141 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 13141; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 275 BP; 92 A; 57 C; 63 G; 63 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||||:||:||||||| |||||:|||| |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              |||||||| ||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 315
AAI48529
ID   AAI48529 standard; DNA; 275 BP.
XX
AC   AAI48529;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17215 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

```
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 17215; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 275 BP; 92 A; 57 C; 63 G; 63 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              |||||||||||||||||:||:|||||||:|||||:|||||||||||||||:|||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||||:||:||||:|||||:||||||||||:|||:||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 316
ABA50362
ID   ABA50362 standard; DNA; 275 BP.
XX
AC   ABA50362;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9057.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 9057; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 275 BP; 92 A; 57 C; 63 G; 63 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              |||||||||||||||||:||:|||||||:|||||:|||||||||||||||:|||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||||:||:||||:|||||:||||||||||:|||:||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 317
AAK42453
ID   AAK42453 standard; DNA; 275 BP.
XX
AC   AAK42453;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17010.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 17010; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 275 BP; 92 A; 57 C; 63 G; 63 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
             ||||||||||||||||||:||:||||||| ||||||:|||||  |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
             ||||||||  ||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
             ||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
             |||:|||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
             :|||:| |||  |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 318
ABS42065
ID   ABS42065 standard; DNA; 275 BP.
XX
AC   ABS42065;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 17055.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 17055; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 275 BP; 92 A; 57 C; 63 G; 63 T; 0 U; 0 Other;

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
             ||||||||||||||||||:||:||||||| ||||||:|||||  |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
             ||||||||  ||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
             ||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
             |||:|||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
             :|||:| |||  |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 319
AAV09125
ID   AAV09125 standard; cDNA; 279 BP.
```

```
XX
AC   AAV09125;
XX
DT   28-MAY-1998  (first entry)
XX
DE   3' nucleotide sequence of clone AJ172_2.
XX
KW   AJ172_2 clone; cDNA library; human adult testes; cell proliferation;
KW   cell differentiation; ss.
XX
OS   Homo sapiens.
XX
PN   US5708157-A.
XX
PD   13-JAN-1998.
XX
PF   26-JUL-1996;   96US-00686878.
XX
PR   26-JUL-1996;   96US-00686878.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Mccoy JM,  Treacy M,  Racie LA,  Merberg D,  Spaulding V,  Jacobs K;
PI   Lavallie ER,  Evans C;
XX
DR   WPI; 1998-109352/10.
XX
PT   Human testes derived cDNA and protein(s) - may be useful for laboratory
PT   use, or for stimulating cell proliferation and differentiation.
XX
PS   Disclosure; Col 81-82; 48pp; English.
XX
CC   This sequence is the 3' nucleotide sequence for the AJ172-2 clone, which
CC   is isolated from a human adult testes cDNA library. When the AJ172_2
CC   sequence was searched in the Genbank database, the clone demonstrated at
CC   least some homology with the following sequences: Friend murine leukemiz
CC   virus; Moloney murine leukemia virus genome; pol protein, Gibbon leukemia
CC   virus; and an EST identified as "yh46a09.s1 Homo sapiens cDNA clone
CC   132760 3'"
XX
SQ   Sequence 279 BP; 97 A; 56 C; 52 G; 61 T; 0 U; 13 Other;

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
             :|||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 320
AAV15136
ID   AAV15136 standard; cDNA; 279 BP.
XX
AC   AAV15136;
XX
DT   29-JUN-1998  (first entry)
XX
DE   Human adult testes cDNA clone AJ172_2 3' cDNA portion.
XX
KW   Human; adult; testes; AJ172_2; secreted protein; cytokine; nutrition;
KW   cell proliferation; differentiation; tumour; immune; ds.
XX
OS   Homo sapiens.
XX
PN   WO9804696-A1.
XX
PD   05-FEB-1998.
XX
PF   23-JUL-1997;   97WO-US012910.
XX
PR   26-JUL-1996;   96US-00686878.
PR   27-SEP-1996;   96US-00721489.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Racie LA,  Merberg D,  Treacy M;
PI   Spaulding V;
XX
DR   WPI; 1998-130687/12.
XX
PT   Polynucleotides and secreted proteins having cytokine and cell
PT   proliferation/differentiation activity - isolated from a human adult
PT   testes cDNA library.
XX
PS   Claim 13; Page 34; 48pp; English.
XX
CC   The present sequence represents the 3' cDNA portion of the clone AJ172_2
CC   from a human adult testes cDNA library. The nucleotide sequences and
CC   protein from clone AJ172_2, have uses in nutrition, cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis regulating activity, chemotactic/chemokinetic
CC   activity, haemostatic or thrombotic activity, receptor/ligand activity,
CC   anti-inflammatory activity, tumour inhibition activity or other
CC   activities. The products can also be used for production of antibodies
CC   and for detection and diagnosis
XX
SQ   Sequence 279 BP; 97 A; 56 C; 52 G; 61 T; 0 U; 13 Other;

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
             :|||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 321
AAV43214
ID   AAV43214 standard; cDNA; 310 BP.
XX
AC   AAV43214;
```

```
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus fragment 3.
XX
KW   Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW   env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2. .310
FT                   /*tag=  a
FT                   /product= "Encodes protein AAW71066"
FT                   /transl_except= (pos:77-79, appears to code for a stop
FT                   codon)
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71066.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Disclosure; Page 182; 286pp; English.
XX
CC   The present sequence represents a multiple sclerosis (MS) associated
CC   retrovirus (MSRV) genomic fragment used in the method of the invention.
CC   The invention provides complete or partial genomic sequences of the MSRV-
CC   1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC   genes. The invention also provides antibodies raised against the
CC   polypeptides. The genomic sequences, polypeptides and antibodies are also
CC   claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC   associated viruses, and also for prevention and treatment of infection
CC   with these viruses
XX
SQ   Sequence 310 BP; 107 A; 78 C; 58 G; 67 T; 0 U; 0 Other;

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             |||||| ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||    | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||  |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             |||||| |||||||| |||||||||| |||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 322
AAX29701
ID   AAX29701 standard; DNA; 310 BP.
XX
AC   AAX29701;
XX
DT   17-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   08-JUN-1999  (first entry)
XX
DE   Clone CL6-5' from MSRV-1.
XX
KW   Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW   rheumatoid polyarthritis; ss.
XX
OS   Viruses.
XX
PN   FR2765588-A1.
XX
PD   08-JAN-1999.
XX
PF   07-JUL-1997;   97FR-00008816.
XX
PR   07-JUL-1997;   97FR-00008816.
XX
PA   (INMR ) BIO MERIEUX.
XX
DR   WPI; 1999-098275/09.
DR   P-PSDB; AAW99551.
XX
PT   Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT   multiple sclerosis or rheumatoid polyarthritis.
XX
PS   Claim 1; Page 36; 83pp; French.
XX
CC   This sequence represents clone CL6-5' from a novel multiple sclerosis
CC   related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC   prophylactic or therapeutic compositions to inhibit expression of a
CC   multiple sclerosis related virus and/or virus associated with rheumatoid
CC   polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC   17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 310 BP; 107 A; 78 C; 58 G; 67 T; 0 U; 0 Other;

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             |||||| ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||    | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||  |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             |||||| |||||||| |||||||||| |||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
```

```
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310
```

RESULT 323
ADG14841
ID ADG14841 standard; cDNA; 310 BP.
XX
AC ADG14841;
XX
DT 26-FEB-2004 (first entry)
XX
DE MSRV associated cDNA #6.
XX
KW ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS Multiple sclerosis associated retrovirus.
XX
PN US2003198647-A1.
XX
PD 23-OCT-2003.
XX
PF 03-APR-2002; 2002US-00114104.
XX
PR 26-NOV-1996; 96US-00756429.
PR 26-NOV-1997; 97US-00979847.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
DR WPI; 2004-032461/03.
XX
PT New isolated nucleic acid and their fragments having the pol gene of a
PT retrovirus, useful for diagnosing, preventing and/or treating multiple
PT sclerosis and/or rheumatoid arthritis.
XX
PS Disclosure; SEQ ID NO 100; 193pp; English.
XX
CC The invention relates to an isolated nucleic acid which comprises the pol
CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC arthritis. The methods and compositions of the present invention are
CC useful for diagnosing, preventing and/or treating multiple sclerosis
CC and/or rheumatoid arthritis. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 310 BP; 107 A; 78 C; 58 G; 67 T; 0 U; 0 Other;

```
  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
              ||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
              ||||||| ||| ||||| || |||||||| |||     | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
              ||||||||| ||||| |||||||||  |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| ||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310
```

RESULT 324
ACH81073/c
ID ACH81073 standard; DNA; 266 BP.
XX
AC ACH81073;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #14268.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 14268; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising

```
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 266 BP; 72 A; 48 C; 83 G; 63 T; 0 U; 0 Other;

  Query Match             3.2%;  Score 236;  DB 1;  Length 266;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  233;  Conservative   19;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       2857 ATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATA 2916
              || |||| ||||||:| ||||||||||||||||||:|||||||||||||||| ||||||
Db        266 ATCCTTCGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCCGGGATA 207

Qy       2917 GYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACA 2976
              |: |||||||||||||||||| ||| |||||||||||||:||:|::|||||||| ||||
Db        206 GCTCCCATCTATTTGGCCAGGCGTTAACCCAAGACTTGAGCCAGTTCTCATACCTTGACA 147

Qy       2977 CTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCC 3036
              ||||||||||||:|||:|||||||||||| |||||:||: ||::||||||||||||||||
Db        146 CTCTTGTCCTTCAGTATGTGGATGATTTGCTTTTAGCCACCCATTCAGAAACCTTGTGCC 87

Qy       3037 ATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAAC 3096
              ||||||| ||||||||:||||| ||||||||||:||||||||||||:|||||||||||
Db          86 ATCAAGCTACCCAAGCGCTCTTGAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAAC 27

Qy       3097 SARARGCTCARCTCTGCTCACAGCAG 3122
                |:|:|||||:|||||||||| |||
Db          26 AAAAGGCTCAGCTCTGCTCACAACAG 1


RESULT 325
AAA98189
ID   AAA98189 standard; DNA; 322 BP.
XX
AC   AAA98189;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human retrovirus LTR DNA fragment HERV-T47D-W5.
XX
KW   Cell-specific expression; tissue-specific expression; gene therapy; LTR;
KW   U3-R segment; long terminal repeat; retroviral expression vector; ds.
XX
OS   Human endogenous retrovirus.
XX
PN   WO200053789-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002064.
XX
PR   10-MAR-1999;   99DE-01010650.
XX
PA   (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI   Leib-Moesch C,  Schoen U,  Baust C;
XX
DR   WPI; 2000-587442/55.
XX
PT   Retroviral expression vector, useful in gene therapy, contains a promoter
PT   from a human endogenous retrovirus to provide cell-specific expression.
XX
PS   Disclosure; Page 58; 67pp; German.
XX
CC   This invention describes a novel retroviral expression vector (A)
CC   containing DNA sequences (I) for packaging vector RNA and for cell-
CC   specific expression of proteins or peptides encoding by heterologous DNA
CC   (II). The sequences controlling cell-specific expression contain a cell-
CC   specifically regulatable promoter region (P) from a human endogenous
CC   retrovirus (HERV) DNA sequence. The invention also describes (a) mRNA and
CC   RNA of (A); (b) prokaryotic and eukaryotic cells containing (A); (c)
CC   eukaryotic cells containing (A) in integrated form; (d) virions
CC   containing a retroviral expression vector RNA derived from (A); (e) a
CC   method for producing the virions of (d); (f) a method for incorporating
CC   protein-encoding nucleic acid sequences into a eukaryotic cell by
CC   infection with the virions of (d); and (g) a retroviral vector system
CC   containing (A) and a packaging cell line, that contains at least one
CC   (recombinant) retrovirus construct that encodes for the packaging
CC   proteins of (A). (A) are used for cell- or tissue-specific expression of
CC   foreign genes for gene therapy and to produce virions for introducing
CC   (II) into the chromosomal DNA of eukaryotic cells, preferably mammalian
CC   and specifically human. (A) retain the advantages of usual retroviral
CC   promoters with all the signal structures required for transcription in a
CC   small region within the U3-R segment, but without their disadvantages
CC   (excessive strength and limited cell specificity). Since (A) are derived
CC   from endogenous (harmless) viral sequences, they do not introduce any new
CC   viral sequences into the genome and recombination will not create new
CC   types of retrovirus. The promoters provide cell or tissue specific
CC   expression, according to which HERV they are derived from
XX
SQ   Sequence 322 BP; 102 A; 76 C; 75 G; 69 T; 0 U; 0 Other;

  Query Match             3.1%;  Score 234.4;  DB 1;  Length 322;
  Best Local Similarity   80.6%;  Pred. No. 0;
  Matches  283;  Conservative    7;  Mismatches   23;  Indels   38;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||||||||||||| |||||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||||||||||||||||||||||||||| ||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||| |||||| ||         |||||||||||||||||||||||||||| || ||||
Db        121 CACACTTGACCAGTCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||| |||||||:||| |:|||||||||| ||||||||||||||||
Db        181 GAGGTAAAGAAATAGACAATCATCTATCACCTGAGAGCACAGTGGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              | |||||||||| |  |||||||| ||||| ||||||||||||||  | ||
Db        241 GCATATAAACCCAGGCATTCGAGCCAGCAACAGCAACCCCCTTTGGGAGCTCT------- 293

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
                                  ||||||||||||||||||||||:||||
Db        294 ----------------------GTTTTCACTCTATTAAATCTTGCAACTGC 322


RESULT 326
ADB84395
ID   ADB84395 standard; DNA; 310 BP.
XX
AC   ADB84395;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 associated DNA sequence #16.
XX
KW   ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Unidentified;.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
```

```
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Claim 31; Page 76; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is a claimed MSRV-associated sequence whose identity cannot be
CC   accurately determined. Note: The SEQ ID numbers for the sequences as
CC   displayed in the main body of the patent do not match the SEQ ID numbers
CC   in the sequence listing. Consequently those sequences mentioned in the
CC   claims may not be the sequences the authors intended to claim.
XX
SQ   Sequence 310 BP; 106 A; 77 C; 58 G; 65 T; 0 U; 4 Other;

  Query Match             3.1%;  Score 233.1;  DB 1;  Length 310;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  262;  Conservative    6;  Mismatches   39;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
              ||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
              ||||||| ||| ||||| || |||||||| |||    | ||| ||||||||  || ||
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCN 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
              ||||||||| ||||| |||||||||| |||||| ||| ||||  ||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAN 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| |||||||||||||:|||||||||| |||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATN 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||||||| |:|||:|||||| |||||||:||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCN 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 327
ADF59598/c
ID   ADF59598 standard; cDNA; 321 BP.
XX
AC   ADF59598;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human polynucleotide sequence SEQ ID NO:1006.
XX
KW   biological activity; genetic engineering; hybridisation probe; oligomer;
KW   primer; chromosome mapping; gene mapping; recombinant protein production;
KW   human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003080795-A2.
XX
PD   02-OCT-2003.
XX
PF   09-AUG-2002; 2002WO-US025485.
XX
PR   09-AUG-2001; 2001US-0311261P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Yang Y,  Wang Z,  Weng G,  Ma Y;
XX
DR   WPI; 2003-876918/81.
DR   P-PSDB; ADF59680.
XX
PT   New polynucleotides, useful as hybridization probes, oligomers or
PT   primers, for chromosome or gene mapping, for the recombinant production
PT   of proteins, and for generating antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 1006; 571pp; English.
XX
CC   The present sequence represents an isolated polynucleotide sequence (I)
CC   from the present invention, which encodes a polypeptide (II) with
CC   biological activity. Also described: (1) a vector comprising (I); (2) an
CC   expression vector comprising (I); (3) a host cell genetically engineered
CC   to comprise (I) which is operatively associated with a regulatory
CC   sequence that modulates expression of (I) in the host cell; (4) a
CC   polypeptide (II) encoded by (I); (5) a composition comprising the
CC   polypeptide of (4) and a carrier; (6) an antibody directed against the
CC   polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
CC   sample; (8) identifying a compound that binds to the polypeptide of (4);
CC   (9) producing the polypeptide of (4); and (10) a collection of
CC   polynucleotides comprising at least one of the polynucleotide sequences
CC   (I). The polynucleotides (I) can be used as hybridisation probes,
CC   oligomers or primers, for chromosome or gene mapping, for the recombinant
CC   production of proteins, and for generating antisense DNA or RNA.
XX
SQ   Sequence 321 BP; 71 A; 76 C; 105 G; 69 T; 0 U; 0 Other;

  Query Match             3.1%;  Score 229.1;  DB 1;  Length 321;
  Best Local Similarity   82.6%;  Pred. No. 0;
  Matches  266;  Conservative    4;  Mismatches   51;  Indels    1;  Gaps    1;

Qy        174 CCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCA 233
              | ||||||| |||||      |||| |||  || |||||| ||| ||||||||||| |||
Db        321 CTACTGCTGATTGCCGTTGTCGCAGACCCACCACTGACTTCCACCCCTCTGGATCCGGCA 262

Qy        234 GGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAA 293
```

```
              |||||| | ||| | | |||| |||||||:|||||||| |||||| |  || |||||| |
Db        261 GGGTGTACACTGCGTTTCTGACCCAGCGAGGCGCCCATCGCCGCTGCGGATCGGGCTAGA 202

Qy        294 GGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACT 353
              |||| |||||||| ||| |||||||||||||  ||||  | | ||| ||||||||||||
Db        201 GGCTCGCCATTGTTCCTTCACGGCTAAGTGCCCAGGTTCATCCCAATCGAGCTGAACACT 142

Qy        354 ANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACT 413
              | |  |||||||||| |||||||||| |||||||||||||||||||:| ||||||||||
Db        141 AGTGGCTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACT 82

Qy        414 TACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAG 473
               |||||||||| |||| |||||||||| |||||| |||||:||| |||| |||||||||
Db         81 CACCACATGGCTCAAGGTTCCATTCCTTAGAATCCATGAGGCCAA-GAACCCCAGGTCAG 23

Qy        474 AGAATACGARGCTTGCCACCAT 495
              |||| |  |:||||||||||||
Db         22 AGAACAAAAGGCTTGCCACCAT 1
```

```
RESULT 328
AAT02816
ID   AAT02816 standard; cDNA; 297 BP.
XX
AC   AAT02816;
XX
DT   20-MAR-1996  (first entry)
XX
DE   Multiple sclerosis associated retrovirus clone MSRV-1B F11-1.
XX
KW   Pol gene; multiple sclerosis; infectious agent; PCR; amplification; ss;
KW   retrovirus; ERV-9; HSERV-9; endogenous; MSRV-2; probe; primer; antibody.
XX
OS   Synthetic.
XX
PN   WO9521256-A1.
XX
PD   10-AUG-1995.
XX
PF   06-FEB-1995;   95WO-FR000142.
XX
PR   04-FEB-1994;   94FR-00001529.
PR   04-FEB-1994;   94FR-00001530.
PR   04-FEB-1994;   94FR-00001531.
PR   04-FEB-1994;   94FR-00001532.
PR   24-NOV-1994;   94FR-00014322.
PR   23-DEC-1994;   94FR-00015810.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Mallet F,  Mandrand B,  Bedin F,  Beseme F;
XX
DR   WPI; 1995-283776/37.
XX
PT   Two pathogenic or infectious agents associated with multiple sclerosis -
PT   one being a retrovirus, useful for early diagnosis, prevention and
PT   treatment.
XX
PS   Claim 3; Page 58; 104pp; French.
XX
CC   The partial nucleotide sequence of the pol gene from the multiple
CC   sclerosis associated infectious agent clone designated MSRV-1B F11-1. The
CC   MSRV-1B (AAT02815-23) and MSRV-2 clones (AAT02824-6) are derived from the
CC   same strain of infectious agent, either POL2 or MS7PG or their variants,
CC   and can be produced by the same cell line, LM7PC or PLI-2 or other cell
CC   culture infected with the multiple sclerosis associated agents. The
CC   clones were isolated by extracting nucleic acids from cells of multiple
CC   sclerosis patients, converting the nucleic acid to cDNA and PCR
CC   amplifying a region of the pol gene conserved in retroviruses. The clones
CC   isolated fell into 2 distinct types: one was similar to the ERV-9 and
CC   HSERV-9 endogenous retroviruses and designated MSRV-1B; the others were
CC   highly homologous to the known agent MSRV-2 and designated MSRV-2B. The
CC   nucleic acid, probes or primers derived from these sequences can be used
CC   for detecting agents associated with multiple sclerosis. The probes may
CC   be used to discriminate between these agents, opt. in conjunction with
CC   amplification primers and antibodies specific for the infectious agents
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              |||| ||||| ||||| | | ||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
              |||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295
```

```
RESULT 329
AAT37974
ID   AAT37974 standard; cDNA; 297 BP.
XX
AC   AAT37974;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   MSRV-1 POL gene clone F11-1.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS   Retroviridae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .297
FT                   /*tag=  a
FT                   /product= "fragment of MSRV-1 polymerase"
FT   misc_feature    277. .297
FT                   /*tag=  b
FT                   /note= "sequence in this region may vary due to
FT                   degenerate probe sequence used to clone this fragment"
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
DR   P-PSDB; AAW12371.
XX
```

```
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 25; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is derived
CC   from a 5' RACE reaction from mRNA isolated from POL2 and MS7PG virions
CC   and forms nucleotides 1-297 of AAT37973. The viral sequences can be used
CC   for the diagnosis, prevention and treatment of RPA caused by infection or
CC   reactivation of the retroviruses. (Updated on 27-AUG-2003 to correct OS
CC   field.)
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| | ,
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 330
AAT96526
ID   AAT96526 standard; cDNA; 297 BP.
XX
AC   AAT96526;
XX
DT   11-MAR-1998  (first entry)
XX
DE   MSRV-1B clone F1-11.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; primer; PCR; ss;
KW   amplification; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .297
FT                   /*tag=  a
FT                   /note= "no stop codon given at end of sequence"
FT   primer_bind     277. .297
FT                   /*tag=  b
FT                   /note= "binds amplification primer AAT96525"
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Example 5; Fig 8; 188pp; French.
XX
CC   This is the nucleotide sequence from clone F1-11, amplified by the
CC   primers AAT96524 and AAT96525 from cDNA synthesised from mRNA associated
CC   with a peak of reverse transcriptase activity from the multiple sclerosis
CC   related virus 1B (MSRV-1B). MSRV-1 sequences are associated with multiple
CC   sclerosis and some are also associated with rheumatoid arthritis. Nucleic
CC   acid sequences from the MSRV-1 virus can be used in the diagnosis of
CC   multiple sclerosis and rheumatoid arthritis
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| | ,
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 331
AAV43116
ID   AAV43116 standard; cDNA; 297 BP.
XX
AC   AAV43116;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus-1 clone F1-11 DNA sequence.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .297
FT                   /*tag=  a
FT                   /product= "clone F1-11 protein"
FT                   /note= "CDS does not contain a stop codon"
XX
```

```
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71050.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 5; Fig 8; 286pp; English.
XX
CC   The present sequence represents the multiple sclerosis (MS) associated
CC   retrovirus-1 (MSRV-1) clone F1-11 DNA sequence. The invention provides
CC   complete or partial genomic sequences of the MSRV-1 pol gene, gag gene
CC   and env gene, and polypeptides encoded by these genes. The invention also
CC   provides antibodies raised against the polypeptides. The genomic
CC   sequences, polypeptides and antibodies are also claimed useful for
CC   diagnosing infection by MS and rheumatoid arthritis-associated viruses,
CC   and also for prevention and treatment of infection with these viruses
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                ||||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 332
ADB84550
ID   ADB84550 standard; cDNA; 297 BP.
XX
AC   ADB84550;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 DNA clone F11-1.
XX
KW   MSRV; ss; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84518.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 5; Fig 8; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV cDNA clone. Note: The SEQ ID numbers for the
CC   sequences as displayed in the main body of the patent do not match the
CC   SEQ ID numbers in the sequence listing. Consequently those sequences
CC   mentioned in the claims may not be the sequences the authors intended to
CC   claim.
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196
```

```
Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 333
ADG14743
ID   ADG14743 standard; cDNA; 297 BP.
XX
AC   ADG14743;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1 clone F-1-11.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 5; SEQ ID NO 2; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 297 BP; 75 A; 77 C; 60 G; 85 T; 0 U; 0 Other;

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 334
ADD26898
ID   ADD26898 standard; DNA; 211 BP.
XX
AC   ADD26898;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human adipocyte Selected Interacting Domain (SID) prey protein DNA #286.
XX
KW   Adipocyte; protein-protein interaction; protein complex;
KW   bait-prey complex; Selected Interacting Domain; SID; drug screening;
KW   drug discovery; metabolic disease; obesity; lipodystrophy;
KW   diabetes mellitus; type 2; non-insulin dependent; NIDDM;
KW   adipogenesis modulation; gene therapy; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200290544-A2.
XX
PD   14-NOV-2002.
XX
PF   03-MAY-2002; 2002WO-EP006333.
XX
PR   04-MAY-2001; 2001US-0288885P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (LYNX-) LYNX THERAPEUTICS INC.
XX
PI   Legrain P,  Whiteside S,  Mao J,  Khrebtukova I,  Luo S;
XX
DR   WPI; 2003-111975/10.
DR   P-PSDB; ADD26900.
XX
PT   New complex of bait and prey between two polypeptides or polynucleotides
PT   encoding the two polypeptides of adipocytes, useful for selecting a
PT   modulating compound that inhibits or activates protein-protein
PT   interactions.
XX
PS   Claim 7; SEQ ID NO 286; 232pp; English.
XX
CC   The invention relates to a bait-prey complex between two adipocyte
CC   polypeptides, or between two polynucleotides encoding adipocyte
CC   polypeptides. The invention also relates to Selected Interacting Domain
CC   (SID) polypeptides which interact with selected bait polypeptides;
CC   polynucleotides encoding SID polypeptides; vectors comprising SID
CC   polynucleotides; recombinant host cells comprising an adipocyte
CC   polynucleotide or a SID-encoding vector; a method of selecting for a
CC   compound which modulates interactions between adipocyte polypeptides;
CC   adipocyte modulator compounds identified using the method; a
CC   pharmaceutical composition comprising an adipocyte modulator, or a SID-
CC   encoding vector or host cell; and a protein chip comprising adipocyte
CC   bait polypeptides. The bait-prey complexes of the invention are useful
CC   for selecting a compound that inhibits or activates protein-protein
CC   interactions between adipocyte polypeptides. The modulatory compounds
CC   identified can be used in the treatment of metabolic diseases such as
CC   obesity, lipodystrophy and type 2 diabetes mellitus, and in the
CC   modulation of adipogenesis. The present sequence represents a
CC   specifically claimed adipocyte SID prey polynucleotide of the invention.
XX
SQ   Sequence 211 BP; 43 A; 74 C; 32 G; 62 T; 0 U; 0 Other;

  Query Match             2.8%;  Score 207.4;  DB 1;  Length 211;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  202;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
```

```
              ||||:||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCA 5791
              |||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCA 211
```

RESULT 335
AAI20303
ID AAI20303 standard; DNA; 278 BP.
XX
AC AAI20303;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #10236 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 10236; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

Query Match 2.7%; Score 205; DB 1; Length 278;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 252; Conservative 0; Mismatches 25; Indels 4; Gaps 3;

```
Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
              ||||||| |||||| |||||| | |||||  | |||||  |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
              |||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
              |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
              |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277
```

RESULT 336
ABA65340
ID ABA65340 standard; DNA; 278 BP.
XX
AC ABA65340;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #13645.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 4; SEQ ID NO 13645; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

Query Match 2.7%; Score 205; DB 1; Length 278;
Best Local Similarity 89.7%; Pred. No. 0;

```
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 337
AAI45507
ID   AAI45507 standard; DNA; 278 BP.
XX
AC   AAI45507;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #14193 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 14193; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;

  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 338
ABA47449
ID   ABA47449 standard; DNA; 278 BP.
XX
AC   ABA47449;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #6144.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 6144; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
```

```
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 339
ABA32438
ID   ABA32438 standard; DNA; 278 BP.
XX
AC   ABA32438;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #10904 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 10904; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 340
AAK39493
ID   AAK39493 standard; DNA; 278 BP.
XX
AC   AAK39493;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 14050.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
```

```
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 14050; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| | ||||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 341
AAK13751
ID   AAK13751 standard; DNA; 278 BP.
XX
AC   AAK13751;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 13742.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 13742; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| | ||||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 342
ABS39084
ID   ABS39084 standard; DNA; 278 BP.
XX
AC   ABS39084;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 14074.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
```

```
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 14074; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             ||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 343
AAI06006
ID   AAI06006 standard; DNA; 278 BP.
XX
AC   AAI06006;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #5997 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 5997; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             ||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 344
ABS13584
ID   ABS13584 standard; DNA; 278 BP.
XX
AC   ABS13584;
XX
DT   19-AUG-2002  (first entry)
XX
```

```
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 13575.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 13575; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 278 BP; 88 A; 68 C; 61 G; 61 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 345
AAI23581/c
ID   AAI23581 standard; DNA; 200 BP.
XX
AC   AAI23581;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13514 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 13514; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
```

```
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 346
ABA68694/c
ID   ABA68694 standard; DNA; 200 BP.
XX
AC   ABA68694;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16999.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 16999; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 347
AAI48896/c
ID   AAI48896 standard; DNA; 200 BP.
XX
AC   AAI48896;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17582 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 17582; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
```

```
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 348
ABA50726/c
ID   ABA50726 standard; DNA; 200 BP.
XX
AC   ABA50726;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9421.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 9421; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 349
ABA35659/c
ID   ABA35659 standard; DNA; 200 BP.
XX
AC   ABA35659;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #14125 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
```

```
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 14125; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 350
AAK42820/c
ID   AAK42820 standard; DNA; 200 BP.
XX
AC   AAK42820;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17377.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 17377; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 351
AAK17036/c
ID   AAK17036 standard; DNA; 200 BP.
XX
AC   AAK17036;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 17027.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
```

```
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 17027; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 352
ABS42449/c
ID   ABS42449 standard; DNA; 200 BP.
XX
AC   ABS42449;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 17439.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 17439; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 353
AAI09197/c
ID   AAI09197 standard; DNA; 200 BP.
XX
AC   AAI09197;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #9188 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
```

```
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 9188; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 354
ABS16875/c
ID   ABS16875 standard; DNA; 200 BP.
XX
AC   ABS16875;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 16866.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 16866; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 200 BP; 44 A; 41 C; 55 G; 60 T; 0 U; 0 Other;

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
```

```
                   |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
                  ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1
```

RESULT 355
AAI56898
ID AAI56898 standard; DNA; 202 BP.
XX
AC AAI56898;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #25584 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 25584; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 202 BP; 64 A; 43 C; 45 G; 50 T; 0 U; 0 Other;

```
  Query Match             2.1%;  Score 156.6;  DB 1;  Length 202;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  183;  Conservative    5;  Mismatches   14;  Indels    2;  Gaps    2;

Qy        973 TAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCC 1032
              |||||||| ||||||| | ||||||||||||||||||||||||||  ||||| ||||| |
Db          1 TAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCCGTTTATGCCCCCT-C 59

Qy       1033 CAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCA 1092
              |||||||||||||||| |||||||||||||||||||||||||||| |:||||::|||| ||
Db         60 CAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCA 119

Qy       1093 GACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              ||||| |||||||| |||| |||| ||||| |||||||||||||:|||||||||||||||
Db        120 GACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAACCCTGATGGCTATATT 179

Qy       1153 GRTGTTTTACAAGGGTTAGGACAA 1176
                  |:|| ||||||||||||||||||
Db        180 GATG-TTTACAAGGGTTAGGACAA 202
```

RESULT 356
AAI46347
ID AAI46347 standard; DNA; 207 BP.
XX
AC AAI46347;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #15033 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 15033; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 207 BP; 67 A; 50 C; 46 G; 44 T; 0 U; 0 Other;

```
  Query Match             2.1%;  Score 156.2;  DB 1;  Length 207;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  188;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       2369 ACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAA 2428
              ||||| |||||||||||||  |||||||||| || ||||||||||||||||||||||||||||
Db          1 ACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGGAATTAATCCTGAA 59

Qy       2429 GTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAA 2488
              |||||||||||||||||||||||||||||| | |||||||||| ||| ||||||||||
Db         60 GTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATCCCATTCAAGTTAA 118

Qy       2489 ACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCA 2548
              |||| ||||||| | ||||||||||||||||||||||| |||| |||||| ||||| |
Db        119 ACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTAGACCCGAGGCCGA 178

Qy       2549 ACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              |||||| ||||||||||||||||||||||
Db        179 TCAAGGACTCCAAAAGATTGTTAAGGAC 206
```

```
RESULT 357
ABS39898
ID   ABS39898 standard; DNA; 207 BP.
XX
AC   ABS39898;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 14888.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 14888; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 207 BP; 67 A; 50 C; 46 G; 44 T; 0 U; 0 Other;

  Query Match             2.1%;  Score 156.2;  DB 1;  Length 207;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  188;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       2369 ACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAA 2428
             ||||| ||||||||||||  ||||||||| || ||||||||||||||||||||||||||||
Db          1 ACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGGAATTAATCCTGAA 59

Qy       2429 GTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAA 2488
             |||||||||||||||||||||||||||||| | |||||||||| |||  ||||||||||
Db         60 GTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATCCCATTCAAGTTAA 118

Qy       2489 ACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCA 2548
             |||| |||||||  | |||||||||||||||||||||||| |||| |||||| ||||| |
Db        119 ACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTAGACCCGAGGCCGA 178

Qy       2549 ACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||||| ||||||||||||||||||||||
Db        179 TCAAGGACTCCAAAAGATTGTTAAGGAC 206


RESULT 358
AAK83421
ID   AAK83421 standard; DNA; 201 BP.
XX
AC   AAK83421;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38233.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
```

PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 38233; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 201 BP; 64 A; 42 C; 45 G; 50 T; 0 U; 0 Other;

Query Match 2.0%; Score 152.6; DB 1; Length 201;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 176; Conservative 3; Mismatches 17; Indels 10; Gaps 1;

```
Qy       7375 AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATY 7434
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||:
Db          6 AGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATC 65

Qy       7435 TATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGC 7494
             ||||||:|||||||  |||  ||||||||||| || ||||||||||||| |  ||| |||
Db         66 TATTGCCTGAGAGCGCAGTGGGAGGGACAATAATTGGGATATAAACCCAGGCATTCAAGC 125

Qy       7495 CGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTA 7554
             |||||| ||| |||| ||||||||||||| ||||||||||||||||||||
Db        126 CGGCAATGGCTACCCTCTTTGGGTCCCCTACCTTTGTATGGGAGCTCTGT---------- 175

Qy       7555 TTTCACTCTATTAAATCTTGCARCTG 7580
              ||  |||||||||||||||| ||:|||
Db        176 TTCTACTCTATTAAATCTTACAACTG 201
```

RESULT 359
ACH84817/c
ID ACH84817 standard; DNA; 169 BP.

```
XX
AC   ACH84817;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #18012.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 18012; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 169 BP; 54 A; 27 C; 49 G; 39 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 145.2;  DB 1;  Length 169;
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  150;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       6879 AAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGG 6938
             ||||||  ||||||||| ||||||||||:|||||||||||||||||||||||||||||||
Db        169 AAACACCAGACCCTGGGACCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGG 110

Qy       6939 ACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGT 6998
             |||||||||  |||||||| || ||||||||||||||||||||||||:||||||| |
Db        109 ACCTCTAGCGACTATAATACTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTTT 50

Qy       6999 TAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              ||| |||||||||||||||| ||||||||:|| ||||||||||
Db         49 TAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTACAAATGG 6


RESULT 360
ACH92789/c
ID   ACH92789 standard; DNA; 194 BP.
XX
AC   ACH92789;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #25984.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 25984; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
```

```
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 194 BP; 51 A; 45 C; 54 G; 44 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 141.9;  DB 1;  Length 194;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  165;  Conservative    3;  Mismatches   26;  Indels    1;  Gaps    1;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||| | ||||||| | |||||| || | |||||||||||||||| |  ||
Db        194 CATTGTTCCTGCGCAGCTAAGTCCTTGGGTTCGTCCCAATTGAGCTGAACACTAGTTGCT 135

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||| |||||||||||||| ||| ||||||||||:| |||||||||| ||| ||
Db        134 GGGTTCCACGGTTCTCTTCTGTGATCCATGGCTTCTAATAGAGCTATAACACTCACCGCA 75

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||| |||||||||||||||||||||||:||| |||| |||| |||| ||| |
Db         74 TGGCCCAAGGTTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGATCAGGGAACAA 16

Qy        481 GARGCTTGCCACCAT 495
               |:||||||  ||||
Db         15 AAGGCTTGCTGCCAT 1


RESULT 361
AAT96481
ID   AAT96481 standard; cDNA; 150 BP.
XX
AC   AAT96481;
XX
DT   04-MAR-1998  (first entry)
XX
DE   Fragment POL2B of the MSRV-1 pol gene sequence.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; RT-PCR; primer; ss;
KW   amplification; reverse transcriptase; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .150
FT                   /*tag=  a
FT                   /product= "fragment of pol protein"
FT                   /note= "no stop codon is given at the end of the
FT                   sequence"
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
DR   P-PSDB; AAW36026.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Claim 16; Fig 28; 188pp; French.
XX
CC   This is the nucleotide sequence of a fragment of the pol gene from the
CC   multiple sclerosis related virus 1 (MSRV-1). The corresponding protein
CC   sequence was used for epitope mapping of the POL2B construct, which
CC   isolated the immunodominant peptides AAW36030-7. The sequences are
CC   associated with multiple sclerosis and some are also associated with
CC   rheumatoid arthritis. Nucleic acid sequences from the MSRV-1 virus can be
CC   used in the diagnosis of multiple sclerosis and rheumatoid arthritis
XX
SQ   Sequence 150 BP; 30 A; 46 C; 27 G; 47 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
              |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
                 |||  ||||||:| |||||||||||||||||:|||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              ||||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 362
AAV43153
ID   AAV43153 standard; cDNA; 150 BP.
XX
AC   AAV43153;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus-1 POL2B encoding DNA sequence.
XX
KW   Multiple sclerosis associated retrovirus-1; MSRV-1; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; POL2B peptide;
KW   ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .150
FT                   /*tag=  a
FT                   /product= "POL2B peptide"
FT                   /note= "CDS does not contain a stop codon"
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
```

```
XX
DR   WPI; 1998-322732/28.
DR   P-PSDB; AAW71051.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 11; Fig 28; 286pp; English.
XX
CC   The present sequence represents a fragment of the multiple sclerosis (MS)
CC   associated retrovirus-1 (MSRV-1) pol DNA sequence. This fragment encodes
CC   a POL2B peptide used in the method of the invention. The invention
CC   provides complete or partial genomic sequences of the MSRV-1 pol gene,
CC   gag gene and env gene, and polypeptides encoded by these genes. The
CC   invention also provides antibodies raised against the polypeptides. The
CC   genomic sequences, polypeptides and antibodies are also claimed useful
CC   for diagnosing infection by MS and rheumatoid arthritis-associated
CC   viruses, and also for prevention and treatment of infection with these
CC   viruses
XX
SQ   Sequence 150 BP; 30 A; 46 C; 27 G; 47 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  |||||||:| |||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 363
ADB84331
ID   ADB84331 standard; DNA; 150 BP.
XX
AC   ADB84331;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1 POL2B gene fragment.
XX
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
DR   P-PSDB; ADB84330.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 11; Fig 28; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 150 BP; 30 A; 46 C; 27 G; 47 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  |||||||:| |||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 364
ADG14777
ID   ADG14777 standard; cDNA; 150 BP.
XX
AC   ADG14777;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1 pol cDNA.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
```

```
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
DR   P-PSDB; ADG14776.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 11; SEQ ID NO 36; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 150 BP; 30 A; 46 C; 27 G; 47 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  |||||:| ||||||||||||||||||:||||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 365
ACH84800
ID   ACH84800 standard; DNA; 174 BP.
XX
AC   ACH84800;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #17995.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 17995; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 174 BP; 24 A; 57 C; 40 G; 53 T; 0 U; 0 Other;

  Query Match             1.8%;  Score 135.6;  DB 1;  Length 174;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  155;  Conservative    1;  Mismatches   16;  Indels    2;  Gaps    1;

Qy        101 TATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTT 160
              |||| |||||||||||||||||||||||||| ||||| |||| ||| ||||||||||||
Db          1 TATTCAATCTTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACAGCTTGAGCTGAGCTTT 60

Qy        161 TGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCC--GCTGACTCCCATC 218
              |||| || |||||||||||||||||||| ||| |||| |||||||  ||||||| |||||
Db         61 CGCTCGCCTTCCACCACTGCTGTTTGCCGCCATCGCAGACCTGCCGTGCTGACTTCCATC 120

Qy        219 CCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATT 272
              ||||| ||||  |||||||||||||||||||| ||||||||||:|||||||||
Db        121 CCTCTAGATCTGGCAGGGTGTCCGCTGTGCTCTTGATCCAGCGAGGCGCCCATT 174


RESULT 366
ADF58618/c
ID   ADF58618 standard; cDNA; 180 BP.
XX
AC   ADF58618;
XX
DT   12-FEB-2004  (first entry)
```

```
XX
DE   Human polynucleotide sequence SEQ ID NO:985.
XX
KW   biological activity; genetic engineering; hybridisation probe; oligomer;
KW   primer; chromosome mapping; gene mapping; recombinant protein production;
KW   human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003080795-A2.
XX
PD   02-OCT-2003.
XX
PF   09-AUG-2002; 2002WO-US025485.
XX
PR   09-AUG-2001; 2001US-0311261P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Yang Y,  Wang Z,  Weng G,  Ma Y;
XX
DR   WPI; 2003-876918/81.
DR   P-PSDB; ADF59659.
XX
PT   New polynucleotides, useful as hybridization probes, oligomers or
PT   primers, for chromosome or gene mapping, for the recombinant production
PT   of proteins, and for generating antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 985; 571pp; English.
XX
CC   The present sequence represents an isolated polynucleotide sequence (I)
CC   from the present invention, which encodes a polypeptide (II) with
CC   biological activity. Also described: (1) a vector comprising (I); (2) an
CC   expression vector comprising (I); (3) a host cell genetically engineered
CC   to comprise (I) which is operatively associated with a regulatory
CC   sequence that modulates expression of (I) in the host cell; (4) a
CC   polypeptide (II) encoded by (I); (5) a composition comprising the
CC   polypeptide of (4) and a carrier; (6) an antibody directed against the
CC   polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
CC   sample; (8) identifying a compound that binds to the polypeptide of (4);
CC   (9) producing the polypeptide of (4); and (10) a collection of
CC   polynucleotides comprising at least one of the polynucleotide sequences
CC   (I). The polynucleotides (I) can be used as hybridisation probes,
CC   oligomers or primers, for chromosome or gene mapping, for the recombinant
CC   production of proteins, and for generating antisense DNA or RNA.
XX
SQ   Sequence 180 BP; 41 A; 39 C; 53 G; 47 T; 0 U; 0 Other;

  Query Match             1.7%;  Score 128.9;  DB 1;  Length 180;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  149;  Conservative    3;  Mismatches   20;  Indels    1;  Gaps    1;

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              ||||| |||||||||||||| |||||||||||||   ||||||||| |||||||| | ||||
Db        180 CTAATCGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCGTGACCCATGACTTC 121

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGS 455
              || ||:| |||||||||| ||||||||||||||| |||||||| |||||||||||||||:
Db        120 TACTAGAGCTATAACACTGACCACATGGCCCAAGGTTCCATTCTTTGGAATCCGTGAGGC 61

Qy        456 CAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||| |||| |||||||||||| |  |:|||||  ||||||||| || ||||
Db         60 CAA-GAACCCCAGGTCAGAGAACAAAAGGCTTGCTGCCATCTTGGGAGTGGCC 9


RESULT 367
ADF58263/c
ID   ADF58263 standard; cDNA; 180 BP.
XX
AC   ADF58263;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human polynucleotide sequence SEQ ID NO:630.
XX
KW   biological activity; genetic engineering; hybridisation probe; oligomer;
KW   primer; chromosome mapping; gene mapping; recombinant protein production;
KW   human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003080795-A2.
XX
PD   02-OCT-2003.
XX
PF   09-AUG-2002; 2002WO-US025485.
XX
PR   09-AUG-2001; 2001US-0311261P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Yang Y,  Wang Z,  Weng G,  Ma Y;
XX
DR   WPI; 2003-876918/81.
DR   P-PSDB; ADF59263.
XX
PT   New polynucleotides, useful as hybridization probes, oligomers or
PT   primers, for chromosome or gene mapping, for the recombinant production
PT   of proteins, and for generating antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 630; 571pp; English.
XX
CC   The present sequence represents an isolated polynucleotide sequence (I)
CC   from the present invention, which encodes a polypeptide (II) with
CC   biological activity. Also described: (1) a vector comprising (I); (2) an
CC   expression vector comprising (I); (3) a host cell genetically engineered
CC   to comprise (I) which is operatively associated with a regulatory
CC   sequence that modulates expression of (I) in the host cell; (4) a
CC   polypeptide (II) encoded by (I); (5) a composition comprising the
CC   polypeptide of (4) and a carrier; (6) an antibody directed against the
CC   polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
CC   sample; (8) identifying a compound that binds to the polypeptide of (4);
CC   (9) producing the polypeptide of (4); and (10) a collection of
CC   polynucleotides comprising at least one of the polynucleotide sequences
CC   (I). The polynucleotides (I) can be used as hybridisation probes,
CC   oligomers or primers, for chromosome or gene mapping, for the recombinant
CC   production of proteins, and for generating antisense DNA or RNA.
XX
SQ   Sequence 180 BP; 41 A; 43 C; 53 G; 43 T; 0 U; 0 Other;

  Query Match             1.7%;  Score 127.3;  DB 1;  Length 180;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  148;  Conservative    3;  Mismatches   21;  Indels    1;  Gaps    1;

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              ||||| |||||||||||||| || |||||||||| |||||||||| |||||||||||||||
Db        180 CTAATCGAGCTGAACACTAGTCGCTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTC 121

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGS 455
              |||||:| |||||||||| ||  |||||||| | |||||||| ||||||||| |||||:
Db        120 TAATAGAGCTATAACACTCACTGCATGGCCCAGGGTTCCATTCTTTGGAATCCATGAGGC 61

Qy        456 CAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||| |||| |||||||||||| |  |:|||||  ||||||||| || ||||
Db         60 CAA-GAACCCCAGGTCAGAGAACAAAAGGCTTGCTGCCATCTTGGGAGTGGCC 9


RESULT 368
AAI23701
ID   AAI23701 standard; DNA; 152 BP.
XX
AC   AAI23701;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13634 for gene expression analysis in human cervical cell sample.
```

```
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 13634; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:|||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 369
ABA68815
ID   ABA68815 standard; DNA; 152 BP.
XX
AC   ABA68815;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #17120.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 17120; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:|||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 370
AAI49015
ID   AAI49015 standard; DNA; 152 BP.
XX
AC   AAI49015;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17701 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
```

```
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 17701; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 371
ABA50844
ID   ABA50844 standard; DNA; 152 BP.
XX
AC   ABA50844;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9539.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 9539; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 372
AAK42939
ID   AAK42939 standard; DNA; 152 BP.
XX
AC   AAK42939;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17496.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
```

```
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 17496; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:|||||||  ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             |||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 373
ABS42569
ID   ABS42569 standard; DNA; 152 BP.
XX
AC   ABS42569;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 17559.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 17559; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:|||||||  ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             |||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 374
AAI09320
ID   AAI09320 standard; DNA; 152 BP.
XX
AC   AAI09320;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #9311 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
```

```
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 9311; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             |||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 375
ABS17000
ID   ABS17000 standard; DNA; 152 BP.
XX
AC   ABS17000;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 16991.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 16991; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 152 BP; 43 A; 22 C; 35 G; 52 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120
```

```
Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143
```

RESULT 376
AAI57411
ID AAI57411 standard; DNA; 133 BP.
XX
AC AAI57411;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #26097 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 26097; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 133 BP; 31 A; 43 C; 23 G; 36 T; 0 U; 0 Other;

Query Match 1.6%; Score 121.8; DB 1; Length 133;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 120; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy       2857 ATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATA 2916
              || |||| ||||||:| |||||||||||| ||||||:||||||||||||||||||||||||
Db          1 ATCCTTCGAACCCAACGTCTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATA 60

Qy       2917 GYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACA 2976
              |:|||||||||||||||||| ||||||||||||||||||:||:|::|||||||||||||||
Db         61 GCCCCCATCTATTTGGCCAGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACA 120

Qy       2977 CTCTTGTCCTTCR 2989
              |||||||||||||:
Db        121 CTCTTGTCCTTCA 133
```

RESULT 377
AAI21346
ID AAI21346 standard; DNA; 155 BP.
XX
AC AAI21346;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #11279 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 11279; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

Query Match 1.6%; Score 121; DB 1; Length 155;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 133; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

```
Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
              ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              || | |||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
              | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155
```

RESULT 378
AAI46625
ID AAI46625 standard; DNA; 155 BP.
XX
AC AAI46625;

```
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #15311 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 15311; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | |||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 379
ABA33486
ID   ABA33486 standard; DNA; 155 BP.
XX
AC   ABA33486;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #11952 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 11952; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | |||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 380
AAK40578
ID   AAK40578 standard; DNA; 155 BP.
XX
AC   AAK40578;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 15135.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 15135; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 381
AAK14842
ID   AAK14842 standard; DNA; 155 BP.
XX
AC   AAK14842;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 14833.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 14833; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 382
ABS40144
ID   ABS40144 standard; DNA; 155 BP.
XX
AC   ABS40144;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 15134.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
```

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 15134; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

Query Match 1.6%; Score 121; DB 1; Length 155;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 133; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

```
Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155
```

RESULT 383
ABS14555
ID ABS14555 standard; DNA; 155 BP.
XX
AC ABS14555;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 14546.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 14546; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 155 BP; 40 A; 49 C; 36 G; 30 T; 0 U; 0 Other;

Query Match 1.6%; Score 121; DB 1; Length 155;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 133; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

```
Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | ||||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
```

```
                 || | ||||||||| |||||||  |||||||  || | ||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
                | ||||||| |||||||| ||| ||||||| ||||
Db         121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 384
AAI28143/c
ID   AAI28143 standard; DNA; 135 BP.
XX
AC   AAI28143;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #18076 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 18076; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| |||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 385
ABA76489/c
ID   ABA76489 standard; DNA; 135 BP.
XX
AC   ABA76489;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #24794.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 24794; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| |||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 386
```

```
AAI57175/c
ID   AAI57175 standard; DNA; 135 BP.
XX
AC   AAI57175;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #25861 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 25861; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 387
ABA41003/c
ID   ABA41003 standard; DNA; 135 BP.
XX
AC   ABA41003;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #19469 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 19469; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 388
AAK51142/c
ID   AAK51142 standard; DNA; 135 BP.
XX
AC   AAK51142;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 25699.
XX
```

```
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 25699; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||| |||||||  |||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 389
AAK25134/c
ID   AAK25134 standard; DNA; 135 BP.
XX
AC   AAK25134;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 25125.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 25125; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||| |||||||  |||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 390
ABS50702/c
ID   ABS50702 standard; DNA; 135 BP.
XX
AC   ABS50702;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 25692.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
```

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 25692; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

Query Match 1.4%; Score 101.8; DB 1; Length 135;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 109; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

```
Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | |||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11
```

RESULT 391
ABS24663/c
ID ABS24663 standard; DNA; 135 BP.
XX
AC ABS24663;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 24654.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 24654; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 135 BP; 31 A; 30 C; 41 G; 33 T; 0 U; 0 Other;

Query Match 1.4%; Score 101.8; DB 1; Length 135;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 109; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

```
Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
```

```
              :||||||| | |||||||| |||||||  |||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 392
ABA72820/c
ID   ABA72820 standard; DNA; 110 BP.
XX
AC   ABA72820;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #21125.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 21125; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
              |||||||||||||||||||||| ||||| |||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 393
AAI53246/c
ID   AAI53246 standard; DNA; 110 BP.
XX
AC   AAI53246;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #21932 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 21932; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
              |||||||||||||||||||||| ||||| |||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 394
ABA38438/c
ID   ABA38438 standard; DNA; 110 BP.
XX
AC   ABA38438;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #16904 for gene expression analysis in human heart cell sample.
XX
```

KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 16904; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

Query Match 1.2%; Score 93; DB 1; Length 110;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 96; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db 110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy 6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
|||| || |||||||||||||||||||| ||:|| ||||||||||
Db 50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6

RESULT 395
AAK47408/c
ID AAK47408 standard; DNA; 110 BP.
XX
AC AAK47408;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 21965.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 21965; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

Query Match 1.2%; Score 93; DB 1; Length 110;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 96; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db 110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy 6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
|||| || |||||||||||||||||||| ||:|| ||||||||||
Db 50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6

RESULT 396
AAK21251/c
ID AAK21251 standard; DNA; 110 BP.
XX
AC AAK21251;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 21242.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.

```
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 21242; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 397
ABS47149/c
ID   ABS47149 standard; DNA; 110 BP.
XX
AC   ABS47149;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 22139.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 22139; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 398
ABS21546/c
ID   ABS21546 standard; DNA; 110 BP.
XX
AC   ABS21546;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 21537.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
```

```
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 21537; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 110 BP; 39 A; 17 C; 26 G; 28 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6

RESULT 399
AAT37982
ID   AAT37982 standard; cDNA; 126 BP.
XX
AC   AAT37982;
XX
DT   27-AUG-2003  (revised)
DT   08-MAY-1997  (first entry)
XX
DE   Rheumatoid polyarthritis associated pol gene clone MSRV1polPR.
XX
KW   Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW   synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW   RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW   diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS   Retroviridae.
XX
PN   EP731168-A1.
XX
PD   11-SEP-1996.
XX
PF   05-MAR-1996;   96EP-00420070.
XX
PR   09-MAR-1995;   95FR-00002960.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Bedin F,  Mandrand B,  Beseme F,  Mallet F;
XX
DR   WPI; 1996-403990/41.
XX
PT   Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT   retroviral or other pathogenic or infectious material from the PLI-2 or
PT   LM7PC cell lines, or related nucleic acid.
XX
PS   Claim 3; Page 35; 68pp; French.
XX
CC   The invention relates to sequences isolated from members of a novel
CC   endogenous retrovirus family associated with sclerotic plaques and
CC   rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC   from cells obtained from sclerotic plaques or synovial joint fluid from
CC   patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC   (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC   the cell lines PLI-2 and MS7PC respectively. This sequence is clone
CC   MSRV1polPR which corresponds to sequences from the pol gene of the
CC   viruses. The sequence is amplified from retroviruses isolated from
CC   synovial joint fluid cells from patients with RPA. The viral sequences
CC   can be used for the diagnosis, prevention and treatment of RPA caused by
CC   infection or reactivation of the retroviruses. (Updated on 27-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 126 BP; 25 A; 36 C; 30 G; 35 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 92;  DB 1;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0;
  Matches   98;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
              :|  |||| || |||||||||||||||: |||||||||||||||| | ||||| ||| ||||
Db          1 GTGCTACCACAGGGGTTCAGGGATAGCTCCCATCTATTTGGCCTGACATTAACCCGAGAC 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTR 3011
              || ||:||:|::|||||| |||||||||||||||| :||| |||||||  | ||  |:
Db         61 TTAAGCCAGTTCTCATACGTGGACACTCTTGTCCTTTGGTACGTGGATGACATCCTGCTG 120

Qy       3012 GCYGCC 3017
              ||: ||
Db        121 GCCTCC 126

RESULT 400
AAI22615/c
ID   AAI22615 standard; DNA; 103 BP.
XX
AC   AAI22615;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #12548 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
```

```
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 12548; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | |||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 401
ABA67692/c
ID   ABA67692 standard; DNA; 103 BP.
XX
AC   ABA67692;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #15997.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 15997; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | |||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 402
AAI47908/c
ID   AAI47908 standard; DNA; 103 BP.
XX
AC   AAI47908;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #16594 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

```
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 16594; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 403
ABA49778/c
ID   ABA49778 standard; DNA; 103 BP.
XX
AC   ABA49778;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #8473.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 8473; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 404
AAK41849/c
ID   AAK41849 standard; DNA; 103 BP.
XX
AC   AAK41849;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 16406.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
```

```
PS   Example 4; SEQ ID NO 16406; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||| |||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 405
AAK16108/c
ID   AAK16108 standard; DNA; 103 BP.
XX
AC   AAK16108;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16099.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 16099; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||| |||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 406
ABS41446/c
ID   ABS41446 standard; DNA; 103 BP.
XX
AC   ABS41446;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 16436.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 16436; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;
```

```
  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||||  |||||||||  ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 407
AAI08291/c
ID   AAI08291 standard; DNA; 103 BP.
XX
AC   AAI08291;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #8282 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 8282; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||||  |||||||||  ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 408
ABS15874/c
ID   ABS15874 standard; DNA; 103 BP.
XX
AC   ABS15874;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 15865.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 15865; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
```

```
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 103 BP; 30 A; 18 C; 24 G; 31 T; 0 U; 0 Other;

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||| | |||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||||  |||||||||  ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 409
AAI20302
ID   AAI20302 standard; DNA; 107 BP.
XX
AC   AAI20302;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #10235 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 10235; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 410
ABA65339
ID   ABA65339 standard; DNA; 107 BP.
XX
AC   ABA65339;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #13644.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 13644; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
```

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 411
AAI45506
ID   AAI45506 standard; DNA; 107 BP.
XX
AC   AAI45506;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #14192 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 14192; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 412
ABA47448
ID   ABA47448 standard; DNA; 107 BP.
XX
AC   ABA47448;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #6143.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 6143; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

```
Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

RESULT 413
ABA32437
ID ABA32437 standard; DNA; 107 BP.
XX
AC ABA32437;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #10903 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 10903; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

Query Match 1.1%; Score 85.2; DB 1; Length 107;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

RESULT 414
AAK39492
ID AAK39492 standard; DNA; 107 BP.
XX
AC AAK39492;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 14049.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 14049; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

Query Match 1.1%; Score 85.2; DB 1; Length 107;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

RESULT 415
AAK13750
ID AAK13750 standard; DNA; 107 BP.
XX
AC AAK13750;
XX
DT 05-NOV-2001 (first entry)

```
XX
DE   Human brain expressed single exon probe SEQ ID NO: 13741.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 13741; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 416
ABS39083
ID   ABS39083 standard; DNA; 107 BP.
XX
AC   ABS39083;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 14073.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 14073; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 417
AAI06005
ID   AAI06005 standard; DNA; 107 BP.
XX
AC   AAI06005;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #5996 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
```

XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 5996; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

Query Match 1.1%; Score 85.2; DB 1; Length 107;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

RESULT 418
ABS13583
ID ABS13583 standard; DNA; 107 BP.
XX
AC ABS13583;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 13574.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.

XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 13574; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 107 BP; 31 A; 34 C; 18 G; 24 T; 0 U; 0 Other;

Query Match 1.1%; Score 85.2; DB 1; Length 107;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60
```

```
Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106
```

RESULT 419
ACH86125/c
ID ACH86125 standard; DNA; 104 BP.
XX
AC ACH86125;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #19320.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 19320; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 104 BP; 24 A; 23 C; 40 G; 17 T; 0 U; 0 Other;

```
  Query Match             1.1%;  Score 85;  DB 1;  Length 104;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches   91;  Conservative    1;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        168 CGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGAT 227
              ||||||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||
Db        104 CGTCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGAT 45

Qy        228 CCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCAT 271
              || |||| || || |||||||||||||||||| ||:||||||||
Db         44 CCGGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCAT 1
```

RESULT 420
AAT37979
ID AAT37979 standard; cDNA; 111 BP.
XX
AC AAT37979;
XX
DT 27-AUG-2003 (revised)
DT 08-MAY-1997 (first entry)
XX
DE MSRV-1B POL gene fragment consensus sequence.
XX
KW Endogenous retrovirus family; sclerotic plaque; rheumatoid polyarthritis;
KW synovial joint fluid; reverse transcriptase polymerase chain reaction;
KW RT-PCR; amplification; primer; exogenous; endogenous; homology; human;
KW diagnosis; prevention; treatment; infection; reactivation; ss.
XX
OS Retroviridae.
XX
PN EP731168-A1.
XX
PD 11-SEP-1996.
XX
PF 05-MAR-1996; 96EP-00420070.
XX
PR 09-MAR-1995; 95FR-00002960.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Perron H, Bedin F, Mandrand B, Beseme F, Mallet F;
XX
DR WPI; 1996-403990/41.
XX
PT Diagnosis, treatment and prevention of rheumatoid polyarthritis - using
PT retroviral or other pathogenic or infectious material from the PLI-2 or
PT LM7PC cell lines, or related nucleic acid.
XX
PS Claim 3; Page 26; 68pp; French.
XX
CC The invention relates to sequences isolated from members of a novel
CC endogenous retrovirus family associated with sclerotic plaques and
CC rheumatoid polyarthritis (RPA). The novel retroviral members are isolated
CC from cells obtained from sclerotic plaques or synovial joint fluid from
CC patients with RPA. Two novel strains of retrovirus are isolated: POL-2
CC (ECACC V92072202) and MS7PG (ECACC V93010816), which are isolated from
CC the cell lines PLI-2 and MS7PC respectively. The sequences AAT37975-9 are
CC consensus sequences derived from the pol gene of those viral sequences
CC designated MSRV-1B, which have 70-88% homology to the human endogenous
CC retrovirus HSERV-9. The sequences were amplified using primers based on
CC reasonably conserved regions of exogenous and endogenous retroviral
CC polymerases and of viruses with reverse transcriptase activity. The
CC sequences can be used for the diagnosis, prevention and treatment of RPA
CC caused by infection or reactivation of the retroviruses. (Updated on 27-

```
CC   AUG-2003 to correct OS field.)
XX
SQ   Sequence 111 BP; 16 A; 17 C; 17 G; 24 T; 0 U; 37 Other;

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 421
AAT96515
ID   AAT96515 standard; cDNA; 111 BP.
XX
AC   AAT96515;
XX
DT   06-MAR-1998  (first entry)
XX
DE   MSRV-1B general consensus sequence from pol region.
XX
KW   Multiple sclerosis related virus 1; env; pol; gag; primer; PCR; ss;
KW   amplification; rheumatoid arthritis; diagnosis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9706260-A1.
XX
PD   20-FEB-1997.
XX
PF   02-AUG-1996;   96WO-FR001244.
XX
PR   03-AUG-1995;   95FR-00009643.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1997-154266/14.
XX
PT   New viral material and nucleotide fragments associated with multiple
PT   sclerosis and rheumatoid arthritis - also related peptide(s) and
PT   antibodies, used for diagnosis, treatment and as vaccines.
XX
PS   Example 1; Page 95; 188pp; French.
XX
CC   This is the sequence of a general consensus region derived from the pol
CC   gene region from the multiple sclerosis related virus 1B (MSRV-1B)
CC   strains POL-2 and MS7PG. The regions were PCR amplified using primers
CC   targetted to the pol gene region. MSRV-1 sequences are associated with
CC   multiple sclerosis and some are also associated with rheumatoid
CC   arthritis. Nucleic acid sequences from the MSRV-1 virus can be used in
CC   the diagnosis of multiple sclerosis and rheumatoid arthritis
XX
SQ   Sequence 111 BP; 16 A; 17 C; 17 G; 24 T; 0 U; 37 Other;

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 422
AAV43121
ID   AAV43121 standard; cDNA; 111 BP.
XX
AC   AAV43121;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Multiple sclerosis associated retrovirus-1B pol DNA sequence.
XX
KW   Multiple sclerosis associated retrovirus-1B; MSRV-1B; MS; pol gene;
KW   gag gene; env gene; rheumatoid arthritis-associated virus; ss.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   WO9823755-A1.
XX
PD   04-JUN-1998.
XX
PF   26-NOV-1997;   97WO-IB001482.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B;
XX
DR   WPI; 1998-322732/28.
XX
PT   New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT   treatment of, e.g. multiple sclerosis.
XX
PS   Example 1; Fig 1; 286pp; English.
XX
CC   The present sequence represents the general consensus multiple sclerosis
CC   (MS) associated retrovirus-1B (MSRV-1B) pol DNA sequence. The invention
CC   provides complete or partial genomic sequences of the MSRV-1 pol gene,
CC   gag gene and env gene, and polypeptides encoded by these genes. The
CC   invention also provides antibodies raised against the polypeptides. The
CC   genomic sequences, polypeptides and antibodies are also claimed useful
CC   for diagnosing infection by MS and rheumatoid arthritis-associated
CC   viruses, and also for prevention and treatment of infection with these
CC   viruses
XX
SQ   Sequence 111 BP; 16 A; 17 C; 17 G; 24 T; 0 U; 37 Other;

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 423
ADB84555
ID   ADB84555 standard; DNA; 111 BP.
XX
AC   ADB84555;
XX
DT   04-DEC-2003  (first entry)
XX
DE   MSRV-1B pol region consensus sequence.
XX
```

```
KW   MSRV; ds; gene; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW   reverse transcriptase; ribonuclease H.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003039664-A1.
XX
PD   27-FEB-2003.
XX
PF   26-NOV-1997;   97US-00979847.
XX
PR   26-NOV-1996;   96US-00756429.
XX
PA   (PERR/) PERRON H.
PA   (BESE/) BESEME F.
PA   (BEDI/) BEDIN F.
PA   (PARA/) PARANHOS-BACCALA G.
PA   (KOMU/) KOMURIAN-PRADEL F.
PA   (JOLI/) JOLIVET-REYNAUD C.
PA   (MAND/) MANDRAND B.
PA   (GARS/) GARSON J A.
PA   (TUKE/) TUKE P W.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2003-512253/48.
XX
PT   New isolated or purified nucleic acid associated with multiple sclerosis
PT   and/or rheumatoid arthritis, useful for detecting a virus associated with
PT   multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
PS   Example 1; Fig 1; 193pp; English.
XX
CC   The invention relates to an isolated or purified nucleic acid from a
CC   virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC   multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC   pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC   proteins or defined peptides (including immunodominant peptides,
CC   antigenic peptides or conserved motifs).Also included are a process for
CC   detecting a virus associated with multiple sclerosis or rheumatoid
CC   arthritis in a biological sample, a nucleic acid probe for the detection
CC   of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC   primer for the amplification by polymerisation of a nucleic acid of a
CC   viral material associated with multiple sclerosis or rheumatoid
CC   arthritis, a polypeptide exhibiting an inhibitory activity on the
CC   proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC   and an antibody directed against the MSRV-1 virus obtained by
CC   immunologically reacting a human or animal body or cells with an
CC   immunogenic agent consisting of the antigenic polypeptide defined above.
CC   The nucleic acids are useful for detecting a biological sample a virus
CC   associated with multiple sclerosis or rheumatoid arthritis, or for
CC   detecting in a biological sample, the presence of or exposure to a virus
CC   associated with multiple sclerosis or rheumatoid arthritis. The present
CC   sequence is an MSRV gene, gene fragment or gene consensus sequence. Note:
CC   The SEQ ID numbers for the sequences as displayed in the main body of the
CC   patent do not match the SEQ ID numbers in the sequence listing.
CC   Consequently those sequences mentioned in the claims may not be the
CC   sequences the authors intended to claim.
XX
SQ   Sequence 111 BP; 16 A; 17 C; 17 G; 24 T; 0 U; 37 Other;

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:||||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110
```

```
RESULT 424
ADG14748
ID   ADG14748 standard; cDNA; 111 BP.
XX
AC   ADG14748;
XX
DT   26-FEB-2004  (first entry)
XX
DE   MSRV-1B consensus sequence.
XX
KW   ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX
OS   Multiple sclerosis associated retrovirus.
XX
PN   US2003198647-A1.
XX
PD   23-OCT-2003.
XX
PF   03-APR-2002; 2002US-00114104.
XX
PR   26-NOV-1996;   96US-00756429.
PR   26-NOV-1997;   97US-00979847.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Perron H,  Beseme F,  Bedin F,  Paranhos-Baccala G;
PI   Komurian-Pradel F,  Jolivet-Reynaud C,  Mandrand B,  Garson JA,  Tuke PW;
XX
DR   WPI; 2004-032461/03.
XX
PT   New isolated nucleic acid and their fragments having the pol gene of a
PT   retrovirus, useful for diagnosing, preventing and/or treating multiple
PT   sclerosis and/or rheumatoid arthritis.
XX
PS   Example 1; SEQ ID NO 7; 193pp; English.
XX
CC   The invention relates to an isolated nucleic acid which comprises the pol
CC   gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC   arthritis. The methods and compositions of the present invention are
CC   useful for diagnosing, preventing and/or treating multiple sclerosis
CC   and/or rheumatoid arthritis. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 111 BP; 16 A; 17 C; 17 G; 24 T; 0 U; 37 Other;

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:||||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy.      2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


Search completed: August  7, 2006, 17:16:04
Job time : 1942 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2006, 16:51:09 ; Search time 2164 Seconds
(without alignments)
3.141 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 7489.6
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 640 seqs, 448195 residues

Total number of hits satisfying chosen parameters: 1278

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Published_Applications_NA_Main

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 638 summaries

Database : pubmaindb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 3549.5 | 47.4 | 4349 | 1 | US-10-450-763-12279 | Sequence 12279, A |
| | 2 | 2890.1 | 38.6 | 2930 | 1 | US-09-902-535-1 | Sequence 1, Appli |
| | 3 | 2889.1 | 38.6 | 2946 | 1 | US-10-114-893-134 | Sequence 134, App |
| | 4 | 2889.1 | 38.6 | 2946 | 1 | US-10-016-249-3 | Sequence 3, Appli |
| | 5 | 2750.9 | 36.7 | 3831 | 1 | US-10-450-763-7531 | Sequence 7531, Ap |
| | 6 | 2706.61 | 36.1 | 2938 | 1 | US-10-632-793-21 | Sequence 21, Appl |
| | 7 | 2706.61 | 36.1 | 2938 | 1 | US-10-717-580-2 | Sequence 2, Appli |
| | 8 | 2626.2 | 35.1 | 3372 | 1 | US-10-632-793-28 | Sequence 28, Appl |
| | 9 | 2626.2 | 35.1 | 3372 | 1 | US-10-717-580-9 | Sequence 9, Appli |
| | 10 | 2551.4 | 34.1 | 2599 | 1 | US-11-028-539-1 | Sequence 1, Appli |
| | 11 | 2315.4 | 30.9 | 2784 | 1 | US-11-028-539-4 | Sequence 4, Appli |
| | 12 | 2256.8 | 30.1 | 2372 | 1 | US-10-632-793-29 | Sequence 29, Appl |
| | 13 | 2256.8 | 30.1 | 2372 | 1 | US-10-717-580-10 | Sequence 10, Appl |
| | 14 | 2205.1 | 29.4 | 2782 | 1 | US-10-632-793-26 | Sequence 26, Appl |
| | 15 | 2205.1 | 29.4 | 2782 | 1 | US-10-717-580-7 | Sequence 7, Appli |
| | 16 | 2201.9 | 29.4 | 2782 | 1 | US-10-133-036-1 | Sequence 1, Appli |
| | 17 | 2141.4 | 28.6 | 2575 | 1 | US-10-717-580-14 | Sequence 14, Appl |
| | 18 | 2036.1 | 27.2 | 2585 | 1 | US-10-717-580-13 | Sequence 13, Appl |
| | 19 | 2032.5 | 27.1 | 2055 | 1 | US-11-028-539-22 | Sequence 22, Appl |
| | 20 | 1851.9 | 24.7 | 2304 | 1 | US-08-979-847-87 | Sequence 87, Appl |
| | 21 | 1851.9 | 24.7 | 2304 | 1 | US-10-114-104-87 | Sequence 87, Appl |
| | 22 | 1848.6 | 24.7 | 2006 | 1 | US-10-632-793-23 | Sequence 23, Appl |
| | 23 | 1848.6 | 24.7 | 2006 | 1 | US-10-717-580-4 | Sequence 4, Appli |
| | 24 | 1812.4 | 24.2 | 2074 | 1 | US-10-416-642-4 | Sequence 4, Appli |
| | 25 | 1780.4 | 23.8 | 2030 | 1 | US-10-637-565-18 | Sequence 18, Appl |
| | 26 | 1607 | 21.5 | 1617 | 1 | US-10-133-036-3 | Sequence 3, Appli |
| | 27 | 1551.5 | 20.7 | 1894 | 1 | US-09-864-761-4444 | Sequence 4444, Ap |
| c | 28 | 1487.1 | 19.9 | 1799 | 1 | US-11-028-539-5 | Sequence 5, Appli |
| | 29 | 1440.3 | 19.2 | 1948 | 1 | US-10-632-793-24 | Sequence 24, Appl |
| | 30 | 1440.3 | 19.2 | 1948 | 1 | US-10-717-580-5 | Sequence 5, Appli |
| | 31 | 1383.7 | 18.5 | 1715 | 1 | US-10-450-763-9505 | Sequence 9505, Ap |
| | 32 | 1329 | 17.7 | 1422 | 1 | US-10-632-793-22 | Sequence 22, Appl |
| | 33 | 1329 | 17.7 | 1422 | 1 | US-10-717-580-3 | Sequence 3, Appli |
| | 34 | 1324.2 | 17.7 | 1481 | 1 | US-08-979-847-105 | Sequence 105, App |
| | 35 | 1324.2 | 17.7 | 1481 | 1 | US-10-114-104-105 | Sequence 105, App |
| | 36 | 1320.4 | 17.6 | 1600 | 1 | US-08-979-847-207 | Sequence 207, App |
| | 37 | 1320.4 | 17.6 | 1600 | 1 | US-10-114-104-207 | Sequence 207, App |
| | 38 | 1309.8 | 17.5 | 1600 | 1 | US-08-979-847-206 | Sequence 206, App |
| | 39 | 1309.8 | 17.5 | 1600 | 1 | US-10-114-104-206 | Sequence 206, App |
| | 40 | 1291.9 | 17.2 | 1597 | 1 | US-08-979-847-205 | Sequence 205, App |
| | 41 | 1291.9 | 17.2 | 1597 | 1 | US-10-114-104-205 | Sequence 205, App |
| | 42 | 1285.7 | 17.2 | 1660 | 1 | US-10-450-763-9412 | Sequence 9412, Ap |
| | 43 | 1285.7 | 17.2 | 1660 | 1 | US-10-450-763-12263 | Sequence 12263, A |
| | 44 | 1285.7 | 17.2 | 1660 | 1 | US-10-450-763-20005 | Sequence 20005, A |
| | 45 | 1270.4 | 17.0 | 1321 | 1 | US-10-632-793-20 | Sequence 20, Appl |
| | 46 | 1270.4 | 17.0 | 1321 | 1 | US-10-717-580-1 | Sequence 1, Appli |
| | 47 | 1229.2 | 16.4 | 1472 | 1 | US-09-864-761-14940 | Sequence 14940, A |
| | 48 | 1201.4 | 16.0 | 1326 | 1 | US-11-028-539-2 | Sequence 2, Appli |
| | 49 | 1117.7 | 14.9 | 1511 | 1 | US-08-979-847-117 | Sequence 117, App |
| | 50 | 1117.7 | 14.9 | 1511 | 1 | US-10-114-104-117 | Sequence 117, App |
| | 51 | 1117.7 | 14.9 | 1511 | 1 | US-10-632-793-1 | Sequence 1, Appli |
| | 52 | 1106.8 | 14.8 | 1136 | 1 | US-10-632-793-25 | Sequence 25, Appl |
| | 53 | 1106.8 | 14.8 | 1136 | 1 | US-10-717-580-6 | Sequence 6, Appli |
| | 54 | 960.2 | 12.8 | 1080 | 1 | US-11-028-539-28 | Sequence 28, Appl |
| | 55 | 952.8 | 12.7 | 1167 | 1 | US-08-979-847-57 | Sequence 57, Appl |
| | 56 | 952.8 | 12.7 | 1167 | 1 | US-10-430-442-61 | Sequence 61, Appl |
| | 57 | 952.8 | 12.7 | 1167 | 1 | US-10-114-104-57 | Sequence 57, Appl |
| | 58 | 948 | 12.7 | 1117 | 1 | US-09-864-761-31445 | Sequence 31445, A |
| | 59 | 948 | 12.7 | 1117 | 1 | US-10-029-386-23846 | Sequence 23846, A |
| | 60 | 936.9 | 12.5 | 1158 | 1 | US-08-979-847-1 | Sequence 1, Appli |
| | 61 | 936.9 | 12.5 | 1158 | 1 | US-10-430-442-1 | Sequence 1, Appli |
| | 62 | 936.9 | 12.5 | 1158 | 1 | US-10-114-104-1 | Sequence 1, Appli |
| c | 63 | 894.401 | 11.9 | 1216 | 1 | US-11-028-539-7 | Sequence 7, Appli |
| c | 64 | 777.2 | 10.4 | 990 | 1 | US-10-301-480-534556 | Sequence 534556, |
| c | 65 | 777.2 | 10.4 | 990 | 1 | US-10-301-480-1147965 | Sequence 1147965, |
| c | 66 | 776.6 | 10.4 | 990 | 1 | US-10-301-480-534555 | Sequence 534555, |
| c | 67 | 776.6 | 10.4 | 990 | 1 | US-10-301-480-1147964 | Sequence 1147964, |
| | 68 | 760.3 | 10.2 | 878 | 1 | US-09-925-065A-918077 | Sequence 918077, |
| | 69 | 760.3 | 10.2 | 878 | 1 | US-09-925-065A-918077 | Sequence 918077, |
| c | 70 | 733.101 | 9.8 | 1000 | 1 | US-10-301-480-567825 | Sequence 567825, |
| c | 71 | 733.101 | 9.8 | 1000 | 1 | US-10-301-480-1181234 | Sequence 1181234, |
| | 72 | 714.8 | 9.5 | 910 | 1 | US-10-450-763-7529 | Sequence 7529, Ap |
| | 73 | 712.9 | 9.5 | 792 | 1 | US-09-864-761-21192 | Sequence 21192, A |
| | 74 | 674.5 | 9.0 | 944 | 1 | US-09-864-761-14911 | Sequence 14911, A |
| | 75 | 662.8 | 8.8 | 924 | 1 | US-10-450-763-8028 | Sequence 8028, Ap |
| c | 76 | 651.8 | 8.7 | 799 | 1 | US-09-925-065A-923664 | Sequence 923664, |
| c | 77 | 651.8 | 8.7 | 799 | 1 | US-09-925-065A-923664 | Sequence 923664, |
| | 78 | 649.4 | 8.7 | 795 | 1 | US-09-925-065A-941867 | Sequence 941867, |
| | 79 | 649.4 | 8.7 | 795 | 1 | US-09-925-065A-941867 | Sequence 941867, |
| c | 80 | 646.8 | 8.6 | 758 | 1 | US-09-925-065A-956235 | Sequence 956235, |
| c | 81 | 646.8 | 8.6 | 758 | 1 | US-09-925-065A-956235 | Sequence 956235, |
| c | 82 | 646 | 8.6 | 902 | 1 | US-10-450-763-7526 | Sequence 7526, Ap |
| | 83 | 637.7 | 8.5 | 841 | 1 | US-10-450-763-12271 | Sequence 12271, A |
| | 84 | 637.2 | 8.5 | 825 | 1 | US-09-925-065A-928581 | Sequence 928581, |
| | 85 | 637.2 | 8.5 | 825 | 1 | US-09-925-065A-928581 | Sequence 928581, |
| | 86 | 637.2 | 8.5 | 825 | 1 | US-09-925-065A-949818 | Sequence 949818, |
| | 87 | 637.2 | 8.5 | 825 | 1 | US-09-925-065A-949818 | Sequence 949818, |
| | 88 | 629.3 | 8.4 | 736 | 1 | US-09-925-065A-918078 | Sequence 918078, |
| | 89 | 629.3 | 8.4 | 736 | 1 | US-09-925-065A-918078 | Sequence 918078, |
| | 90 | 623.6 | 8.3 | 791 | 1 | US-09-925-065A-925162 | Sequence 925162, |
| | 91 | 623.6 | 8.3 | 791 | 1 | US-09-925-065A-925162 | Sequence 925162, |
| | 92 | 621.1 | 8.3 | 719 | 1 | US-09-925-065A-941866 | Sequence 941866, |
| | 93 | 621.1 | 8.3 | 719 | 1 | US-09-925-065A-941866 | Sequence 941866, |
| c | 94 | 619.1 | 8.3 | 762 | 1 | US-09-925-065A-936448 | Sequence 936448, |
| c | 95 | 619.1 | 8.3 | 762 | 1 | US-09-925-065A-936448 | Sequence 936448, |
| c | 96 | 593.7 | 7.9 | 729 | 1 | US-09-925-065A-937950 | Sequence 937950, |
| c | 97 | 593.7 | 7.9 | 729 | 1 | US-09-925-065A-937950 | Sequence 937950, |
| c | 98 | 591.7 | 7.9 | 727 | 1 | US-09-925-065A-937711 | Sequence 937711, |
| c | 99 | 591.7 | 7.9 | 727 | 1 | US-09-925-065A-937711 | Sequence 937711, |
| | 100 | 590.8 | 7.9 | 745 | 1 | US-09-925-065A-954498 | Sequence 954498, |
| | 101 | 590.8 | 7.9 | 745 | 1 | US-09-925-065A-954498 | Sequence 954498, |
| c | 102 | 589.7 | 7.9 | 725 | 1 | US-09-925-065A-926262 | Sequence 926262, |
| c | 103 | 589.7 | 7.9 | 725 | 1 | US-09-925-065A-926262 | Sequence 926262, |
| c | 104 | 586.3 | 7.8 | 720 | 1 | US-09-925-065A-937967 | Sequence 937967, |
| c | 105 | 586.3 | 7.8 | 720 | 1 | US-09-925-065A-937967 | Sequence 937967, |
| | 106 | 586 | 7.8 | 741 | 1 | US-08-979-847-9 | Sequence 9, Appli |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 107 | 586 | 7.8 | 741 | 1 | US-10-430-442-9 | Sequence 9, Appli |
| | 108 | 586 | 7.8 | 741 | 1 | US-10-114-104-9 | Sequence 9, Appli |
| c | 109 | 585.9 | 7.8 | 709 | 1 | US-09-925-065A-936449 | Sequence 936449, |
| c | 110 | 585.9 | 7.8 | 709 | 1 | US-09-925-065A-936449 | Sequence 936449, |
| | 111 | 582.1 | 7.8 | 600 | 1 | US-10-717-580-47 | Sequence 47, Appl |
| c | 112 | 581.5 | 7.8 | 720 | 1 | US-09-925-065A-937161 | Sequence 937161, |
| c | 113 | 581.5 | 7.8 | 720 | 1 | US-09-925-065A-937161 | Sequence 937161, |
| c | 114 | 579.7 | 7.7 | 803 | 1 | US-09-925-065A-940856 | Sequence 940856, |
| c | 115 | 579.7 | 7.7 | 803 | 1 | US-09-925-065A-940856 | Sequence 940856, |
| | 116 | 564.8 | 7.5 | 758 | 1 | US-08-979-847-112 | Sequence 112, App |
| | 117 | 564.8 | 7.5 | 758 | 1 | US-10-114-104-112 | Sequence 112, App |
| c | 118 | 553.3 | 7.4 | 632 | 1 | US-09-925-065A-734747 | Sequence 734747, |
| c | 119 | 553.3 | 7.4 | 632 | 1 | US-09-925-065A-734747 | Sequence 734747, |
| | 120 | 553.2 | 7.4 | 752 | 1 | US-09-925-065A-956040 | Sequence 956040, |
| | 121 | 553.2 | 7.4 | 752 | 1 | US-09-925-065A-956040 | Sequence 956040, |
| | 122 | 551.6 | 7.4 | 752 | 1 | US-09-925-065A-956039 | Sequence 956039, |
| | 123 | 551.6 | 7.4 | 752 | 1 | US-09-925-065A-956039 | Sequence 956039, |
| c | 124 | 551.2 | 7.4 | 751 | 1 | US-09-925-065A-948058 | Sequence 948058, |
| c | 125 | 551.2 | 7.4 | 751 | 1 | US-09-925-065A-948058 | Sequence 948058, |
| c | 126 | 550.5 | 7.4 | 626 | 1 | US-09-925-065A-799227 | Sequence 799227, |
| c | 127 | 550.5 | 7.4 | 626 | 1 | US-09-925-065A-799227 | Sequence 799227, |
| c | 128 | 535.5 | 7.1 | 663 | 1 | US-10-301-480-299445 | Sequence 299445, |
| c | 129 | 535.5 | 7.1 | 663 | 1 | US-10-301-480-912854 | Sequence 912854, |
| c | 130 | 532.7 | 7.1 | 636 | 1 | US-09-925-065A-359698 | Sequence 359698, |
| c | 131 | 532.7 | 7.1 | 636 | 1 | US-09-925-065A-359698 | Sequence 359698, |
| c | 132 | 532.7 | 7.1 | 636 | 1 | US-09-925-065A-359699 | Sequence 359699, |
| c | 133 | 532.7 | 7.1 | 636 | 1 | US-09-925-065A-359699 | Sequence 359699, |
| c | 134 | 532.7 | 7.1 | 637 | 1 | US-10-301-480-430862 | Sequence 430862, |
| c | 135 | 532.7 | 7.1 | 637 | 1 | US-10-301-480-430863 | Sequence 430863, |
| c | 136 | 532.7 | 7.1 | 637 | 1 | US-10-301-480-1044271 | Sequence 1044271, |
| c | 137 | 532.7 | 7.1 | 637 | 1 | US-10-301-480-1044272 | Sequence 1044272, |
| c | 138 | 532.6 | 7.1 | 735 | 1 | US-09-925-065A-926128 | Sequence 926128, |
| c | 139 | 532.6 | 7.1 | 735 | 1 | US-09-925-065A-926128 | Sequence 926128, |
| | 140 | 523.9 | 7.0 | 645 | 1 | US-08-979-847-8 | Sequence 8, Appli |
| | 141 | 523.9 | 7.0 | 645 | 1 | US-10-430-442-8 | Sequence 8, Appli |
| | 142 | 523.9 | 7.0 | 645 | 1 | US-10-114-104-8 | Sequence 8, Appli |
| | 143 | 522.6 | 7.0 | 530 | 1 | US-10-717-580-48 | Sequence 48, Appl |
| | 144 | 522.6 | 7.0 | 719 | 1 | US-09-925-065A-928579 | Sequence 928579, |
| | 145 | 522.6 | 7.0 | 719 | 1 | US-09-925-065A-928579 | Sequence 928579, |
| | 146 | 521.4 | 7.0 | 719 | 1 | US-09-925-065A-928577 | Sequence 928577, |
| | 147 | 521.4 | 7.0 | 719 | 1 | US-09-925-065A-928577 | Sequence 928577, |
| c | 148 | 520.3 | 6.9 | 591 | 1 | US-10-029-386-22647 | Sequence 22647, A |
| | 149 | 519.3 | 6.9 | 629 | 1 | US-09-925-065A-935057 | Sequence 935057, |
| | 150 | 519.3 | 6.9 | 629 | 1 | US-09-925-065A-935057 | Sequence 935057, |
| | 151 | 517.4 | 6.9 | 635 | 1 | US-08-979-847-102 | Sequence 102, App |
| | 152 | 517.4 | 6.9 | 635 | 1 | US-10-114-104-102 | Sequence 102, App |
| | 153 | 516.8 | 6.9 | 591 | 1 | US-10-632-793-19 | Sequence 19, Appl |
| | 154 | 516.8 | 6.9 | 591 | 1 | US-10-717-580-31 | Sequence 31, Appl |
| | 155 | 513.6 | 6.9 | 598 | 1 | US-09-925-065A-919228 | Sequence 919228, |
| | 156 | 513.6 | 6.9 | 598 | 1 | US-09-925-065A-919228 | Sequence 919228, |
| | 157 | 512.8 | 6.8 | 651 | 1 | US-09-925-065A-212368 | Sequence 212368, |
| | 158 | 512.8 | 6.8 | 651 | 1 | US-09-925-065A-212368 | Sequence 212368, |
| | 159 | 510.9 | 6.8 | 619 | 1 | US-09-925-065A-896106 | Sequence 896106, |
| | 160 | 510.9 | 6.8 | 619 | 1 | US-09-925-065A-896106 | Sequence 896106, |
| c | 161 | 510.6 | 6.8 | 656 | 1 | US-09-925-065A-943242 | Sequence 943242, |
| c | 162 | 510.6 | 6.8 | 656 | 1 | US-09-925-065A-943242 | Sequence 943242, |
| | 163 | 510.6 | 6.8 | 657 | 1 | US-09-925-065A-923210 | Sequence 923210, |
| | 164 | 510.6 | 6.8 | 657 | 1 | US-09-925-065A-923210 | Sequence 923210, |
| | 165 | 510.3 | 6.8 | 542 | 1 | US-10-040-916-48 | Sequence 48, Appl |
| c | 166 | 508 | 6.8 | 626 | 1 | US-09-925-065A-920958 | Sequence 920958, |
| c | 167 | 508 | 6.8 | 626 | 1 | US-09-925-065A-920958 | Sequence 920958, |
| c | 168 | 507.1 | 6.8 | 627 | 1 | US-09-925-065A-210294 | Sequence 210294, |
| c | 169 | 507.1 | 6.8 | 627 | 1 | US-09-925-065A-210294 | Sequence 210294, |
| c | 170 | 507.1 | 6.8 | 627 | 1 | US-09-925-065A-210296 | Sequence 210296, |
| c | 171 | 507.1 | 6.8 | 627 | 1 | US-09-925-065A-210296 | Sequence 210296, |
| c | 172 | 507.1 | 6.8 | 628 | 1 | US-10-301-480-297639 | Sequence 297639, |
| c | 173 | 507.1 | 6.8 | 628 | 1 | US-10-301-480-297641 | Sequence 297641, |
| c | 174 | 507.1 | 6.8 | 628 | 1 | US-10-301-480-911048 | Sequence 911048, |
| c | 175 | 507.1 | 6.8 | 628 | 1 | US-10-301-480-911050 | Sequence 911050, |
| | 176 | 505.2 | 6.7 | 678 | 1 | US-10-632-793-18 | Sequence 18, Appl |
| | 177 | 505.2 | 6.7 | 678 | 1 | US-10-717-580-29 | Sequence 29, Appl |
| | 178 | 502.7 | 6.7 | 628 | 1 | US-09-925-065A-770753 | Sequence 770753, |
| | 179 | 502.7 | 6.7 | 628 | 1 | US-09-925-065A-770753 | Sequence 770753, |
| | 180 | 502.3 | 6.7 | 628 | 1 | US-09-925-065A-835279 | Sequence 835279, |
| | 181 | 502.3 | 6.7 | 628 | 1 | US-09-925-065A-835279 | Sequence 835279, |
| c | 182 | 501.7 | 6.7 | 627 | 1 | US-09-925-065A-760795 | Sequence 760795, |
| c | 183 | 501.7 | 6.7 | 627 | 1 | US-09-925-065A-760795 | Sequence 760795, |
| c | 184 | 501.3 | 6.7 | 627 | 1 | US-09-925-065A-829558 | Sequence 829558, |
| c | 185 | 501.3 | 6.7 | 627 | 1 | US-09-925-065A-829558 | Sequence 829558, |
| | 186 | 500.1 | 6.7 | 607 | 1 | US-09-925-065A-944021 | Sequence 944021, |
| | 187 | 500.1 | 6.7 | 607 | 1 | US-09-925-065A-944021 | Sequence 944021, |
| c | 188 | 498.4 | 6.7 | 620 | 1 | US-09-925-065A-933320 | Sequence 933320, |
| c | 189 | 498.4 | 6.7 | 620 | 1 | US-09-925-065A-933320 | Sequence 933320, |
| | 190 | 496 | 6.6 | 608 | 1 | US-09-925-065A-923906 | Sequence 923906, |
| | 191 | 496 | 6.6 | 608 | 1 | US-09-925-065A-923906 | Sequence 923906, |
| c | 192 | 495.1 | 6.6 | 618 | 1 | US-10-301-480-441252 | Sequence 441252, |
| c | 193 | 495.1 | 6.6 | 618 | 1 | US-10-301-480-1054661 | Sequence 1054661, |
| | 194 | 494.6 | 6.6 | 609 | 1 | US-10-301-480-475703 | Sequence 475703, |
| | 195 | 494.6 | 6.6 | 609 | 1 | US-10-301-480-1089112 | Sequence 1089112, |
| | 196 | 494.2 | 6.6 | 609 | 1 | US-10-301-480-475704 | Sequence 475704, |
| | 197 | 494.2 | 6.6 | 609 | 1 | US-10-301-480-1089113 | Sequence 1089113, |
| c | 198 | 491.7 | 6.6 | 613 | 1 | US-09-925-065A-371249 | Sequence 371249, |
| c | 199 | 491.7 | 6.6 | 613 | 1 | US-09-925-065A-371249 | Sequence 371249, |
| | 200 | 490.2 | 6.5 | 622 | 1 | US-09-925-065A-946840 | Sequence 946840, |
| | 201 | 490.2 | 6.5 | 622 | 1 | US-09-925-065A-946840 | Sequence 946840, |
| c | 202 | 490 | 6.5 | 618 | 1 | US-10-301-480-384873 | Sequence 384873, |
| c | 203 | 490 | 6.5 | 618 | 1 | US-10-301-480-998282 | Sequence 998282, |
| | 204 | 489.8 | 6.5 | 583 | 1 | US-10-450-763-24195 | Sequence 24195, A |
| c | 205 | 489.6 | 6.5 | 614 | 1 | US-09-925-065A-309261 | Sequence 309261, |
| c | 206 | 489.6 | 6.5 | 614 | 1 | US-09-925-065A-309261 | Sequence 309261, |
| c | 207 | 488.8 | 6.5 | 618 | 1 | US-10-301-480-384874 | Sequence 384874, |
| c | 208 | 488.8 | 6.5 | 618 | 1 | US-10-301-480-998283 | Sequence 998283, |
| c | 209 | 488.4 | 6.5 | 614 | 1 | US-09-925-065A-309262 | Sequence 309262, |
| c | 210 | 488.4 | 6.5 | 614 | 1 | US-09-925-065A-309262 | Sequence 309262, |
| c | 211 | 487 | 6.5 | 566 | 1 | US-10-301-480-228755 | Sequence 228755, |
| c | 212 | 487 | 6.5 | 566 | 1 | US-10-301-480-228756 | Sequence 228756, |
| c | 213 | 487 | 6.5 | 566 | 1 | US-10-301-480-842164 | Sequence 842164, |
| c | 214 | 487 | 6.5 | 566 | 1 | US-10-301-480-842165 | Sequence 842165, |
| c | 215 | 486.2 | 6.5 | 595 | 1 | US-09-925-065A-195740 | Sequence 195740, |
| c | 216 | 486.2 | 6.5 | 595 | 1 | US-09-925-065A-195740 | Sequence 195740, |
| c | 217 | 486.2 | 6.5 | 595 | 1 | US-10-301-480-284876 | Sequence 284876, |
| c | 218 | 486.2 | 6.5 | 595 | 1 | US-10-301-480-898285 | Sequence 898285, |
| | 219 | 486.2 | 6.5 | 681 | 1 | US-09-925-065A-928578 | Sequence 928578, |
| | 220 | 486.2 | 6.5 | 681 | 1 | US-09-925-065A-928578 | Sequence 928578, |
| c | 221 | 486 | 6.5 | 570 | 1 | US-09-925-065A-132544 | Sequence 132544, |
| c | 222 | 486 | 6.5 | 570 | 1 | US-09-925-065A-132544 | Sequence 132544, |
| c | 223 | 486 | 6.5 | 570 | 1 | US-09-925-065A-132545 | Sequence 132545, |
| c | 224 | 486 | 6.5 | 570 | 1 | US-09-925-065A-132545 | Sequence 132545, |
| c | 225 | 484.8 | 6.5 | 615 | 1 | US-09-925-065A-955763 | Sequence 955763, |
| c | 226 | 484.8 | 6.5 | 615 | 1 | US-09-925-065A-955763 | Sequence 955763, |
| c | 227 | 483.8 | 6.5 | 614 | 1 | US-09-925-065A-955838 | Sequence 955838, |
| c | 228 | 483.8 | 6.5 | 614 | 1 | US-09-925-065A-955838 | Sequence 955838, |
| | 229 | 483.4 | 6.5 | 593 | 1 | US-09-925-065A-409227 | Sequence 409227, |
| | 230 | 483.4 | 6.5 | 593 | 1 | US-09-925-065A-409227 | Sequence 409227, |
| c | 231 | 483.3 | 6.5 | 583 | 1 | US-10-301-480-373537 | Sequence 373537, |
| c | 232 | 483.3 | 6.5 | 583 | 1 | US-10-301-480-986946 | Sequence 986946, |
| | 233 | 483 | 6.4 | 593 | 1 | US-09-925-065A-409228 | Sequence 409228, |
| | 234 | 483 | 6.4 | 593 | 1 | US-09-925-065A-409228 | Sequence 409228, |
| c | 235 | 481.9 | 6.4 | 631 | 1 | US-09-925-065A-755821 | Sequence 755821, |
| c | 236 | 481.9 | 6.4 | 631 | 1 | US-09-925-065A-755821 | Sequence 755821, |
| c | 237 | 481.5 | 6.4 | 620 | 1 | US-09-925-065A-948057 | Sequence 948057, |
| c | 238 | 481.5 | 6.4 | 620 | 1 | US-09-925-065A-948057 | Sequence 948057, |
| c | 239 | 481.1 | 6.4 | 549 | 1 | US-09-864-761-7607 | Sequence 7607, Ap |
| c | 240 | 480.9 | 6.4 | 630 | 1 | US-09-925-065A-739053 | Sequence 739053, |
| c | 241 | 480.9 | 6.4 | 630 | 1 | US-09-925-065A-739053 | Sequence 739053, |
| c | 242 | 480.8 | 6.4 | 604 | 1 | US-10-301-480-203869 | Sequence 203869, |
| c | 243 | 480.8 | 6.4 | 604 | 1 | US-10-301-480-817278 | Sequence 817278, |
| | 244 | 479.9 | 6.4 | 612 | 1 | US-10-301-480-382968 | Sequence 382968, |
| | 245 | 479.9 | 6.4 | 612 | 1 | US-10-301-480-996377 | Sequence 996377, |
| c | 246 | 479.9 | 6.4 | 644 | 1 | US-09-925-065A-925326 | Sequence 925326, |
| c | 247 | 479.9 | 6.4 | 644 | 1 | US-09-925-065A-925326 | Sequence 925326, |
| c | 248 | 478.9 | 6.4 | 659 | 1 | US-09-925-065A-936791 | Sequence 936791, |
| c | 249 | 478.9 | 6.4 | 659 | 1 | US-09-925-065A-936791 | Sequence 936791, |
| c | 250 | 478.9 | 6.4 | 659 | 1 | US-09-925-065A-936792 | Sequence 936792, |
| c | 251 | 478.9 | 6.4 | 659 | 1 | US-09-925-065A-936792 | Sequence 936792, |
| c | 252 | 477.3 | 6.4 | 559 | 1 | US-10-029-386-8924 | Sequence 8924, Ap |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 253 | 476.2 | 6.4 | 656 | 1 | US-09-925-065A-769980 | Sequence 769980, |
| | 254 | 476.2 | 6.4 | 656 | 1 | US-09-925-065A-769980 | Sequence 769980, |
| c | 255 | 475.2 | 6.3 | 586 | 1 | US-09-925-065A-785421 | Sequence 785421, |
| c | 256 | 475.2 | 6.3 | 586 | 1 | US-09-925-065A-785421 | Sequence 785421, |
| c | 257 | 474.8 | 6.3 | 597 | 1 | US-09-925-065A-103629 | Sequence 103629, |
| c | 258 | 474.8 | 6.3 | 597 | 1 | US-09-925-065A-103629 | Sequence 103629, |
| c | 259 | 474 | 6.3 | 614 | 1 | US-09-925-065A-764810 | Sequence 764810, |
| c | 260 | 474 | 6.3 | 614 | 1 | US-09-925-065A-764810 | Sequence 764810, |
| c | 261 | 473.9 | 6.3 | 568 | 1 | US-10-029-386-563 | Sequence 563, App |
| c | 262 | 473.6 | 6.3 | 586 | 1 | US-09-925-065A-785419 | Sequence 785419, |
| c | 263 | 473.6 | 6.3 | 586 | 1 | US-09-925-065A-785419 | Sequence 785419, |
| c | 264 | 473.6 | 6.3 | 586 | 1 | US-09-925-065A-785420 | Sequence 785420, |
| c | 265 | 473.6 | 6.3 | 586 | 1 | US-09-925-065A-785420 | Sequence 785420, |
| c | 266 | 473.3 | 6.3 | 573 | 1 | US-09-925-065A-296721 | Sequence 296721, |
| c | 267 | 473.3 | 6.3 | 573 | 1 | US-09-925-065A-296721 | Sequence 296721, |
| | 268 | 471 | 6.3 | 633 | 1 | US-09-925-065A-633131 | Sequence 633131, |
| | 269 | 471 | 6.3 | 633 | 1 | US-09-925-065A-633131 | Sequence 633131, |
| c | 270 | 470.9 | 6.3 | 617 | 1 | US-09-925-065A-926542 | Sequence 926542, |
| c | 271 | 470.9 | 6.3 | 617 | 1 | US-09-925-065A-926542 | Sequence 926542, |
| | 272 | 470.8 | 6.3 | 623 | 1 | US-10-301-480-280971 | Sequence 280971, |
| | 273 | 470.8 | 6.3 | 623 | 1 | US-10-301-480-894380 | Sequence 894380, |
| c | 274 | 470.2 | 6.3 | 560 | 1 | US-09-925-065A-926543 | Sequence 926543, |
| c | 275 | 470.2 | 6.3 | 560 | 1 | US-09-925-065A-926543 | Sequence 926543, |
| | 276 | 469.8 | 6.3 | 600 | 1 | US-09-925-065A-191356 | Sequence 191356, |
| | 277 | 469.8 | 6.3 | 600 | 1 | US-09-925-065A-191356 | Sequence 191356, |
| c | 278 | 469 | 6.3 | 560 | 1 | US-09-925-065A-948890 | Sequence 948890, |
| c | 279 | 469 | 6.3 | 560 | 1 | US-09-925-065A-948890 | Sequence 948890, |
| | 280 | 468.9 | 6.3 | 601 | 1 | US-09-925-065A-307095 | Sequence 307095, |
| | 281 | 468.9 | 6.3 | 601 | 1 | US-09-925-065A-307095 | Sequence 307095, |
| c | 282 | 468.6 | 6.3 | 560 | 1 | US-09-925-065A-926544 | Sequence 926544, |
| c | 283 | 468.6 | 6.3 | 560 | 1 | US-09-925-065A-926544 | Sequence 926544, |
| | 284 | 468.4 | 6.3 | 603 | 1 | US-09-925-065A-205418 | Sequence 205418, |
| | 285 | 468.4 | 6.3 | 603 | 1 | US-09-925-065A-205418 | Sequence 205418, |
| | 286 | 468.1 | 6.3 | 486 | 1 | US-10-717-580-49 | Sequence 49, Appl |
| | 287 | 467.7 | 6.2 | 588 | 1 | US-09-925-065A-419320 | Sequence 419320, |
| | 288 | 467.7 | 6.2 | 588 | 1 | US-09-925-065A-419320 | Sequence 419320, |
| c | 289 | 467.3 | 6.2 | 552 | 1 | US-09-925-065A-925327 | Sequence 925327, |
| c | 290 | 467.3 | 6.2 | 552 | 1 | US-09-925-065A-925327 | Sequence 925327, |
| | 291 | 467.2 | 6.2 | 603 | 1 | US-09-925-065A-205417 | Sequence 205417, |
| | 292 | 467.2 | 6.2 | 603 | 1 | US-09-925-065A-205417 | Sequence 205417, |
| c | 293 | 466.9 | 6.2 | 613 | 1 | US-10-027-632-213458 | Sequence 213458, |
| c | 294 | 466.9 | 6.2 | 613 | 1 | US-10-027-632-213458 | Sequence 213458, |
| | 295 | 466.9 | 6.2 | 614 | 1 | US-09-925-065A-775883 | Sequence 775883, |
| | 296 | 466.9 | 6.2 | 614 | 1 | US-09-925-065A-775883 | Sequence 775883, |
| c | 297 | 466.1 | 6.2 | 552 | 1 | US-09-925-065A-948281 | Sequence 948281, |
| c | 298 | 466.1 | 6.2 | 552 | 1 | US-09-925-065A-948281 | Sequence 948281, |
| | 299 | 466.1 | 6.2 | 588 | 1 | US-09-925-065A-419319 | Sequence 419319, |
| | 300 | 466.1 | 6.2 | 588 | 1 | US-09-925-065A-419319 | Sequence 419319, |
| c | 301 | 465.7 | 6.2 | 552 | 1 | US-09-925-065A-925328 | Sequence 925328, |
| c | 302 | 465.7 | 6.2 | 552 | 1 | US-09-925-065A-925328 | Sequence 925328, |
| | 303 | 465 | 6.2 | 568 | 1 | US-09-925-065A-411946 | Sequence 411946, |
| | 304 | 465 | 6.2 | 568 | 1 | US-09-925-065A-411946 | Sequence 411946, |
| | 305 | 464.8 | 6.2 | 599 | 1 | US-10-301-480-293356 | Sequence 293356, |
| | 306 | 464.8 | 6.2 | 599 | 1 | US-10-301-480-906765 | Sequence 906765, |
| c | 307 | 464.7 | 6.2 | 591 | 1 | US-10-301-480-251972 | Sequence 251972, |
| c | 308 | 464.7 | 6.2 | 591 | 1 | US-10-301-480-251974 | Sequence 251974, |
| c | 309 | 464.7 | 6.2 | 591 | 1 | US-10-301-480-865381 | Sequence 865381, |
| c | 310 | 464.7 | 6.2 | 591 | 1 | US-10-301-480-865383 | Sequence 865383, |
| | 311 | 463.6 | 6.2 | 599 | 1 | US-10-301-480-293355 | Sequence 293355, |
| | 312 | 463.6 | 6.2 | 599 | 1 | US-10-301-480-906764 | Sequence 906764, |
| | 313 | 461.4 | 6.2 | 633 | 1 | US-09-925-065A-758566 | Sequence 758566, |
| | 314 | 461.4 | 6.2 | 633 | 1 | US-09-925-065A-758566 | Sequence 758566, |
| | 315 | 460.7 | 6.2 | 568 | 1 | US-09-864-761-16537 | Sequence 16537, A |
| c | 316 | 458.7 | 6.1 | 585 | 1 | US-09-925-065A-159043 | Sequence 159043, |
| c | 317 | 458.7 | 6.1 | 585 | 1 | US-09-925-065A-159043 | Sequence 159043, |
| c | 318 | 458.7 | 6.1 | 585 | 1 | US-09-925-065A-159045 | Sequence 159045, |
| c | 319 | 458.7 | 6.1 | 585 | 1 | US-09-925-065A-159045 | Sequence 159045, |
| | 320 | 455.7 | 6.1 | 569 | 1 | US-09-925-065A-295923 | Sequence 295923, |
| | 321 | 455.7 | 6.1 | 569 | 1 | US-09-925-065A-295923 | Sequence 295923, |
| | 322 | 454.1 | 6.1 | 569 | 1 | US-09-925-065A-295924 | Sequence 295924, |
| | 323 | 454.1 | 6.1 | 569 | 1 | US-09-925-065A-295924 | Sequence 295924, |
| | 324 | 454.1 | 6.1 | 569 | 1 | US-09-925-065A-295925 | Sequence 295925, |
| | 325 | 454.1 | 6.1 | 569 | 1 | US-09-925-065A-295925 | Sequence 295925, |
| c | 326 | 454 | 6.1 | 548 | 1 | US-09-864-761-13065 | Sequence 13065, A |
| | 327 | 453.6 | 6.1 | 606 | 1 | US-09-925-065A-421237 | Sequence 421237, |
| | 328 | 453.6 | 6.1 | 606 | 1 | US-09-925-065A-421237 | Sequence 421237, |
| | 329 | 453.6 | 6.1 | 606 | 1 | US-09-925-065A-421238 | Sequence 421238, |
| | 330 | 453.6 | 6.1 | 606 | 1 | US-09-925-065A-421238 | Sequence 421238, |
| c | 331 | 452.6 | 6.0 | 606 | 1 | US-10-301-480-486080 | Sequence 486080, |
| c | 332 | 452.6 | 6.0 | 606 | 1 | US-10-301-480-486081 | Sequence 486081, |
| c | 333 | 452.6 | 6.0 | 606 | 1 | US-10-301-480-1099489 | Sequence 1099489, |
| c | 334 | 452.6 | 6.0 | 606 | 1 | US-10-301-480-1099490 | Sequence 1099490, |
| c | 335 | 451.8 | 6.0 | 564 | 1 | US-10-301-480-416983 | Sequence 416983, |
| c | 336 | 451.8 | 6.0 | 564 | 1 | US-10-301-480-1030392 | Sequence 1030392, |
| | 337 | 450.7 | 6.0 | 564 | 1 | US-10-301-480-372764 | Sequence 372764, |
| | 338 | 450.7 | 6.0 | 564 | 1 | US-10-301-480-986173 | Sequence 986173, |
| | 339 | 450.4 | 6.0 | 608 | 1 | US-09-925-065A-950020 | Sequence 950020, |
| | 340 | 450.4 | 6.0 | 608 | 1 | US-09-925-065A-950020 | Sequence 950020, |
| | 341 | 449.1 | 6.0 | 564 | 1 | US-10-301-480-372765 | Sequence 372765, |
| | 342 | 449.1 | 6.0 | 564 | 1 | US-10-301-480-372766 | Sequence 372766, |
| | 343 | 449.1 | 6.0 | 564 | 1 | US-10-301-480-986174 | Sequence 986174, |
| | 344 | 449.1 | 6.0 | 564 | 1 | US-10-301-480-986175 | Sequence 986175, |
| c | 345 | 448.8 | 6.0 | 561 | 1 | US-09-925-065A-344446 | Sequence 344446, |
| c | 346 | 448.8 | 6.0 | 561 | 1 | US-09-925-065A-344446 | Sequence 344446, |
| c | 347 | 448.8 | 6.0 | 565 | 1 | US-09-925-065A-790020 | Sequence 790020, |
| c | 348 | 448.8 | 6.0 | 565 | 1 | US-09-925-065A-790020 | Sequence 790020, |
| | 349 | 448.6 | 6.0 | 578 | 1 | US-10-029-386-10274 | Sequence 10274, A |
| c | 350 | 448.2 | 6.0 | 581 | 1 | US-09-925-065A-787519 | Sequence 787519, |
| c | 351 | 448.2 | 6.0 | 581 | 1 | US-09-925-065A-787519 | Sequence 787519, |
| c | 352 | 448 | 6.0 | 570 | 1 | US-09-925-065A-948681 | Sequence 948681, |
| c | 353 | 448 | 6.0 | 570 | 1 | US-09-925-065A-948681 | Sequence 948681, |
| c | 354 | 446.9 | 6.0 | 577 | 1 | US-09-925-065A-945278 | Sequence 945278, |
| c | 355 | 446.9 | 6.0 | 577 | 1 | US-09-925-065A-945278 | Sequence 945278, |
| c | 356 | 446.8 | 6.0 | 560 | 1 | US-09-925-065A-789246 | Sequence 789246, |
| c | 357 | 446.8 | 6.0 | 560 | 1 | US-09-925-065A-789246 | Sequence 789246, |
| c | 358 | 446.8 | 6.0 | 562 | 1 | US-09-925-065A-789955 | Sequence 789955, |
| c | 359 | 446.8 | 6.0 | 562 | 1 | US-09-925-065A-789955 | Sequence 789955, |
| c | 360 | 445.4 | 5.9 | 558 | 1 | US-09-925-065A-756176 | Sequence 756176, |
| c | 361 | 445.4 | 5.9 | 558 | 1 | US-09-925-065A-756176 | Sequence 756176, |
| c | 362 | 445.1 | 5.9 | 559 | 1 | US-09-864-761-7501 | Sequence 7501, Ap |
| | 363 | 439.6 | 5.9 | 598 | 1 | US-10-027-632-192839 | Sequence 192839, |
| | 364 | 439.6 | 5.9 | 598 | 1 | US-10-027-632-192839 | Sequence 192839, |
| | 365 | 437.8 | 5.8 | 548 | 1 | US-09-925-065A-919226 | Sequence 919226, |
| | 366 | 437.8 | 5.8 | 548 | 1 | US-09-925-065A-919226 | Sequence 919226, |
| | 367 | 433.2 | 5.8 | 491 | 1 | US-09-864-761-32803 | Sequence 32803, A |
| c | 368 | 430.3 | 5.7 | 551 | 1 | US-10-301-480-473577 | Sequence 473577, |
| c | 369 | 430.3 | 5.7 | 551 | 1 | US-10-301-480-1086986 | Sequence 1086986, |
| | 370 | 429.5 | 5.7 | 566 | 1 | US-10-029-386-20359 | Sequence 20359, A |
| c | 371 | 429.3 | 5.7 | 523 | 1 | US-09-925-065A-296722 | Sequence 296722, |
| c | 372 | 429.3 | 5.7 | 523 | 1 | US-09-925-065A-296722 | Sequence 296722, |
| c | 373 | 427.7 | 5.7 | 532 | 1 | US-09-925-065A-210295 | Sequence 210295, |
| c | 374 | 427.7 | 5.7 | 532 | 1 | US-09-925-065A-210295 | Sequence 210295, |
| | 375 | 426.9 | 5.7 | 558 | 1 | US-09-925-065A-857674 | Sequence 857674, |
| | 376 | 426.9 | 5.7 | 558 | 1 | US-09-925-065A-857674 | Sequence 857674, |
| | 377 | 426.8 | 5.7 | 582 | 1 | US-10-301-480-49432 | Sequence 49432, A |
| | 378 | 426.8 | 5.7 | 582 | 1 | US-10-301-480-662841 | Sequence 662841, |
| c | 379 | 425.7 | 5.7 | 530 | 1 | US-10-301-480-297640 | Sequence 297640, |
| c | 380 | 425.7 | 5.7 | 530 | 1 | US-10-301-480-911049 | Sequence 911049, |
| c | 381 | 425.2 | 5.7 | 560 | 1 | US-10-301-480-473605 | Sequence 473605, |
| c | 382 | 425.2 | 5.7 | 560 | 1 | US-10-301-480-1087014 | Sequence 1087014, |
| c | 383 | 424.6 | 5.7 | 563 | 1 | US-10-301-480-414023 | Sequence 414023, |
| c | 384 | 424.6 | 5.7 | 563 | 1 | US-10-301-480-414024 | Sequence 414024, |
| c | 385 | 424.6 | 5.7 | 563 | 1 | US-10-301-480-1027432 | Sequence 1027432, |
| c | 386 | 424.6 | 5.7 | 563 | 1 | US-10-301-480-1027433 | Sequence 1027433, |
| | 387 | 423 | 5.6 | 507 | 1 | US-10-029-386-6530 | Sequence 6530, Ap |
| c | 388 | 423 | 5.6 | 563 | 1 | US-10-301-480-414025 | Sequence 414025, |
| c | 389 | 423 | 5.6 | 563 | 1 | US-10-301-480-414026 | Sequence 414026, |
| c | 390 | 423 | 5.6 | 563 | 1 | US-10-301-480-1027434 | Sequence 1027434, |
| c | 391 | 423 | 5.6 | 563 | 1 | US-10-301-480-1027435 | Sequence 1027435, |
| c | 392 | 422 | 5.6 | 551 | 1 | US-09-925-065A-406936 | Sequence 406936, |
| c | 393 | 422 | 5.6 | 551 | 1 | US-09-925-065A-406936 | Sequence 406936, |
| c | 394 | 421.3 | 5.6 | 572 | 1 | US-09-925-065A-210293 | Sequence 210293, |
| c | 395 | 421.3 | 5.6 | 572 | 1 | US-09-925-065A-210293 | Sequence 210293, |
| | 396 | 421.1 | 5.6 | 549 | 1 | US-09-925-065A-952783 | Sequence 952783, |
| | 397 | 421.1 | 5.6 | 549 | 1 | US-09-925-065A-952783 | Sequence 952783, |
| | 398 | 420.9 | 5.6 | 558 | 1 | US-09-864-761-16507 | Sequence 16507, A |

| | | | | | | |
|---|---|---|---|---|---|---|
| c 399 | 420.2 | 5.6 | 521 | 1 | US-10-029-386-4312 | Sequence 4312, Ap |
| c 400 | 420.1 | 5.6 | 548 | 1 | US-09-925-065A-803434 | Sequence 803434, |
| c 401 | 420.1 | 5.6 | 548 | 1 | US-09-925-065A-803434 | Sequence 803434, |
| 402 | 420.1 | 5.6 | 548 | 1 | US-09-925-065A-956897 | Sequence 956897, |
| 403 | 420.1 | 5.6 | 548 | 1 | US-09-925-065A-956897 | Sequence 956897, |
| c 404 | 419.9 | 5.6 | 565 | 1 | US-09-925-065A-932905 | Sequence 932905, |
| c 405 | 419.9 | 5.6 | 565 | 1 | US-09-925-065A-932905 | Sequence 932905, |
| c 406 | 419.3 | 5.6 | 537 | 1 | US-09-925-065A-406907 | Sequence 406907, |
| c 407 | 419.3 | 5.6 | 537 | 1 | US-09-925-065A-406907 | Sequence 406907, |
| c 408 | 416.9 | 5.6 | 566 | 1 | US-10-301-480-297638 | Sequence 297638, |
| c 409 | 416.9 | 5.6 | 566 | 1 | US-10-301-480-911047 | Sequence 911047, |
| 410 | 409.6 | 5.5 | 488 | 1 | US-09-864-761-1777 | Sequence 1777, Ap |
| 411 | 409.4 | 5.5 | 564 | 1 | US-09-925-065A-924272 | Sequence 924272, |
| 412 | 409.4 | 5.5 | 564 | 1 | US-09-925-065A-924272 | Sequence 924272, |
| 413 | 404.1 | 5.4 | 552 | 1 | US-09-925-065A-926278 | Sequence 926278, |
| 414 | 404.1 | 5.4 | 552 | 1 | US-09-925-065A-926278 | Sequence 926278, |
| 415 | 403 | 5.4 | 425 | 1 | US-10-717-580-45 | Sequence 45, Appl |
| 416 | 403 | 5.4 | 427 | 1 | US-10-717-580-46 | Sequence 46, Appl |
| 417 | 402.5 | 5.4 | 552 | 1 | US-09-925-065A-948763 | Sequence 948763, |
| 418 | 402.5 | 5.4 | 552 | 1 | US-09-925-065A-948763 | Sequence 948763, |
| c 419 | 402 | 5.4 | 534 | 1 | US-09-925-065A-341103 | Sequence 341103, |
| c 420 | 402 | 5.4 | 534 | 1 | US-09-925-065A-341103 | Sequence 341103, |
| c 421 | 402 | 5.4 | 534 | 1 | US-09-925-065A-341104 | Sequence 341104, |
| c 422 | 402 | 5.4 | 534 | 1 | US-09-925-065A-341104 | Sequence 341104, |
| c 423 | 401.6 | 5.4 | 518 | 1 | US-10-301-480-473604 | Sequence 473604, |
| c 424 | 401.6 | 5.4 | 518 | 1 | US-10-301-480-1087013 | Sequence 1087013, |
| c 425 | 401.3 | 5.4 | 537 | 1 | US-10-301-480-455987 | Sequence 455987, |
| c 426 | 401.3 | 5.4 | 537 | 1 | US-10-301-480-1069396 | Sequence 1069396, |
| c 427 | 400.4 | 5.3 | 504 | 1 | US-09-864-761-7027 | Sequence 7027, Ap |
| c 428 | 400.4 | 5.3 | 534 | 1 | US-09-925-065A-341105 | Sequence 341105, |
| c 429 | 400.4 | 5.3 | 534 | 1 | US-09-925-065A-341105 | Sequence 341105, |
| c 430 | 400.4 | 5.3 | 534 | 1 | US-09-925-065A-341106 | Sequence 341106, |
| c 431 | 400.4 | 5.3 | 534 | 1 | US-09-925-065A-341106 | Sequence 341106, |
| 432 | 399.3 | 5.3 | 526 | 1 | US-10-301-480-319631 | Sequence 319631, |
| 433 | 399.3 | 5.3 | 526 | 1 | US-10-301-480-933040 | Sequence 933040, |
| 434 | 396.9 | 5.3 | 523 | 1 | US-09-925-065A-237209 | Sequence 237209, |
| 435 | 396.9 | 5.3 | 523 | 1 | US-09-925-065A-237209 | Sequence 237209, |
| c 436 | 393.2 | 5.2 | 508 | 1 | US-09-925-065A-406935 | Sequence 406935, |
| c 437 | 393.2 | 5.2 | 508 | 1 | US-09-925-065A-406935 | Sequence 406935, |
| 438 | 390.9 | 5.2 | 488 | 1 | US-10-301-480-452763 | Sequence 452763, |
| 439 | 390.9 | 5.2 | 488 | 1 | US-10-301-480-1066172 | Sequence 1066172, |
| 440 | 389.3 | 5.2 | 488 | 1 | US-10-301-480-452764 | Sequence 452764, |
| 441 | 389.3 | 5.2 | 488 | 1 | US-10-301-480-452765 | Sequence 452765, |
| 442 | 389.3 | 5.2 | 488 | 1 | US-10-301-480-1066173 | Sequence 1066173, |
| 443 | 389.3 | 5.2 | 488 | 1 | US-10-301-480-1066174 | Sequence 1066174, |
| 444 | 389.2 | 5.2 | 434 | 1 | US-09-925-065A-567704 | Sequence 567704, |
| 445 | 389.2 | 5.2 | 434 | 1 | US-09-925-065A-567704 | Sequence 567704, |
| c 446 | 389.2 | 5.2 | 437 | 1 | US-09-925-065A-295343 | Sequence 295343, |
| c 447 | 389.2 | 5.2 | 437 | 1 | US-09-925-065A-295343 | Sequence 295343, |
| 448 | 387.5 | 5.2 | 474 | 1 | US-10-029-386-23983 | Sequence 23983, A |
| 449 | 386.4 | 5.2 | 495 | 1 | US-09-925-065A-919227 | Sequence 919227, |
| 450 | 386.4 | 5.2 | 495 | 1 | US-09-925-065A-919227 | Sequence 919227, |
| 451 | 386.4 | 5.2 | 495 | 1 | US-09-925-065A-945490 | Sequence 945490, |
| 452 | 386.4 | 5.2 | 495 | 1 | US-09-925-065A-945490 | Sequence 945490, |
| c 453 | 386.3 | 5.2 | 485 | 1 | US-10-301-480-452762 | Sequence 452762, |
| c 454 | 386.3 | 5.2 | 485 | 1 | US-10-301-480-1066171 | Sequence 1066171, |
| c 455 | 385.8 | 5.2 | 432 | 1 | US-10-301-480-372204 | Sequence 372204, |
| c 456 | 385.8 | 5.2 | 432 | 1 | US-10-301-480-985613 | Sequence 985613, |
| c 457 | 385.4 | 5.1 | 519 | 1 | US-09-925-065A-387464 | Sequence 387464, |
| c 458 | 385.4 | 5.1 | 519 | 1 | US-09-925-065A-387464 | Sequence 387464, |
| c 459 | 384.8 | 5.1 | 472 | 1 | US-09-925-065A-383962 | Sequence 383962, |
| c 460 | 384.8 | 5.1 | 472 | 1 | US-09-925-065A-383962 | Sequence 383962, |
| c 461 | 384.8 | 5.1 | 472 | 1 | US-09-925-065A-383963 | Sequence 383963, |
| c 462 | 384.8 | 5.1 | 472 | 1 | US-09-925-065A-383963 | Sequence 383963, |
| c 463 | 383.2 | 5.1 | 472 | 1 | US-09-925-065A-383964 | Sequence 383964, |
| c 464 | 383.2 | 5.1 | 472 | 1 | US-09-925-065A-383964 | Sequence 383964, |
| c 465 | 383.2 | 5.1 | 472 | 1 | US-09-925-065A-383965 | Sequence 383965, |
| c 466 | 383.2 | 5.1 | 472 | 1 | US-09-925-065A-383965 | Sequence 383965, |
| 467 | 382.5 | 5.1 | 438 | 1 | US-08-979-847-201 | Sequence 201, App |
| 468 | 382.5 | 5.1 | 438 | 1 | US-10-114-104-201 | Sequence 201, App |
| 469 | 381.3 | 5.1 | 438 | 1 | US-08-979-847-128 | Sequence 128, App |
| 470 | 381.3 | 5.1 | 438 | 1 | US-08-979-847-129 | Sequence 129, App |
| 471 | 381.3 | 5.1 | 438 | 1 | US-08-979-847-131 | Sequence 131, App |
| 472 | 381.3 | 5.1 | 438 | 1 | US-10-114-104-128 | Sequence 128, App |
| 473 | 381.3 | 5.1 | 438 | 1 | US-10-114-104-129 | Sequence 129, App |
| 474 | 381.3 | 5.1 | 438 | 1 | US-10-114-104-131 | Sequence 131, App |
| c 475 | 379.8 | 5.1 | 490 | 1 | US-09-925-065A-159044 | Sequence 159044, |
| c 476 | 379.8 | 5.1 | 490 | 1 | US-09-925-065A-159044 | Sequence 159044, |
| 477 | 379.7 | 5.1 | 438 | 1 | US-08-979-847-130 | Sequence 130, App |
| 478 | 379.7 | 5.1 | 438 | 1 | US-10-114-104-130 | Sequence 130, App |
| 479 | 378.6 | 5.1 | 433 | 1 | US-09-864-761-16318 | Sequence 16318, A |
| 480 | 377.7 | 5.0 | 422 | 1 | US-09-864-761-13967 | Sequence 13967, A |
| c 481 | 372.8 | 5.0 | 483 | 1 | US-10-301-480-251973 | Sequence 251973, |
| c 482 | 372.8 | 5.0 | 483 | 1 | US-10-301-480-865382 | Sequence 865382, |
| c 483 | 367.8 | 4.9 | 471 | 1 | US-09-925-065A-251122 | Sequence 251122, |
| c 484 | 367.8 | 4.9 | 471 | 1 | US-09-925-065A-251122 | Sequence 251122, |
| c 485 | 365.7 | 4.9 | 497 | 1 | US-10-301-480-467319 | Sequence 467319, |
| c 486 | 365.7 | 4.9 | 497 | 1 | US-10-301-480-1080728 | Sequence 1080728, |
| c 487 | 365.7 | 4.9 | 500 | 1 | US-09-925-065A-399976 | Sequence 399976, |
| c 488 | 365.7 | 4.9 | 500 | 1 | US-09-925-065A-399976 | Sequence 399976, |
| 489 | 361.2 | 4.8 | 454 | 1 | US-10-301-480-337623 | Sequence 337623, |
| 490 | 361.2 | 4.8 | 454 | 1 | US-10-301-480-951032 | Sequence 951032, |
| 491 | 359.6 | 4.8 | 454 | 1 | US-10-301-480-337622 | Sequence 337622, |
| 492 | 359.6 | 4.8 | 454 | 1 | US-10-301-480-951031 | Sequence 951031, |
| 493 | 359.2 | 4.8 | 451 | 1 | US-09-925-065A-259058 | Sequence 259058, |
| 494 | 359.2 | 4.8 | 451 | 1 | US-09-925-065A-259058 | Sequence 259058, |
| 495 | 357.6 | 4.8 | 451 | 1 | US-09-925-065A-259057 | Sequence 259057, |
| 496 | 357.6 | 4.8 | 451 | 1 | US-09-925-065A-259057 | Sequence 259057, |
| c 497 | 357.2 | 4.8 | 448 | 1 | US-10-450-763-23371 | Sequence 23371, A |
| c 498 | 357 | 4.8 | 457 | 1 | US-10-301-480-330802 | Sequence 330802, |
| c 499 | 357 | 4.8 | 457 | 1 | US-10-301-480-944211 | Sequence 944211, |
| c 500 | 354.5 | 4.7 | 440 | 1 | US-09-925-065A-225538 | Sequence 225538, |
| c 501 | 354.5 | 4.7 | 440 | 1 | US-09-925-065A-225538 | Sequence 225538, |
| c 502 | 353.6 | 4.7 | 439 | 1 | US-09-764-881-183 | Sequence 183, App |
| c 503 | 353.6 | 4.7 | 439 | 1 | US-10-242-747-183 | Sequence 183, App |
| c 504 | 353.6 | 4.7 | 439 | 1 | US-10-158-057-365 | Sequence 365, App |
| 505 | 353.5 | 4.7 | 439 | 1 | US-10-301-480-310136 | Sequence 310136, |
| 506 | 353.5 | 4.7 | 439 | 1 | US-10-301-480-923545 | Sequence 923545, |
| 507 | 345 | 4.6 | 429 | 1 | US-08-979-847-138 | Sequence 138, App |
| 508 | 345 | 4.6 | 429 | 1 | US-10-114-104-138 | Sequence 138, App |
| 509 | 343 | 4.6 | 481 | 1 | US-10-301-480-337624 | Sequence 337624, |
| 510 | 343 | 4.6 | 481 | 1 | US-10-301-480-951033 | Sequence 951033, |
| 511 | 343 | 4.6 | 483 | 1 | US-09-925-065A-259059 | Sequence 259059, |
| 512 | 343 | 4.6 | 483 | 1 | US-09-925-065A-259059 | Sequence 259059, |
| 513 | 341 | 4.6 | 429 | 1 | US-08-979-847-136 | Sequence 136, App |
| 514 | 341 | 4.6 | 429 | 1 | US-10-114-104-136 | Sequence 136, App |
| c 515 | 340.7 | 4.5 | 453 | 1 | US-10-301-480-277403 | Sequence 277403, |
| c 516 | 340.7 | 4.5 | 453 | 1 | US-10-301-480-890812 | Sequence 890812, |
| 517 | 339.4 | 4.5 | 429 | 1 | US-08-979-847-135 | Sequence 135, App |
| 518 | 339.4 | 4.5 | 429 | 1 | US-08-979-847-137 | Sequence 137, App |
| 519 | 339.4 | 4.5 | 429 | 1 | US-10-114-104-135 | Sequence 135, App |
| 520 | 339.4 | 4.5 | 429 | 1 | US-10-114-104-137 | Sequence 137, App |
| c 521 | 339.1 | 4.5 | 453 | 1 | US-10-301-480-277404 | Sequence 277404, |
| c 522 | 339.1 | 4.5 | 453 | 1 | US-10-301-480-890813 | Sequence 890813, |
| c 523 | 339 | 4.5 | 410 | 1 | US-09-880-107-538 | Sequence 538, App |
| c 524 | 339 | 4.5 | 410 | 1 | US-09-960-706-167 | Sequence 167, App |
| c 525 | 339 | 4.5 | 410 | 1 | US-09-873-319-105 | Sequence 105, App |
| 526 | 338.6 | 4.5 | 397 | 1 | US-09-925-065A-925163 | Sequence 925163, |
| 527 | 338.6 | 4.5 | 397 | 1 | US-09-925-065A-925163 | Sequence 925163, |
| c 528 | 335.8 | 4.5 | 405 | 1 | US-09-864-761-3364 | Sequence 3364, Ap |
| c 529 | 332.3 | 4.4 | 390 | 1 | US-09-925-065A-954031 | Sequence 954031, |
| c 530 | 332.3 | 4.4 | 390 | 1 | US-09-925-065A-954031 | Sequence 954031, |
| c 531 | 330.3 | 4.4 | 388 | 1 | US-09-925-065A-948754 | Sequence 948754, |
| c 532 | 330.3 | 4.4 | 388 | 1 | US-09-925-065A-948754 | Sequence 948754, |
| c 533 | 329.3 | 4.4 | 440 | 1 | US-09-925-065A-187359 | Sequence 187359, |
| c 534 | 329.3 | 4.4 | 440 | 1 | US-09-925-065A-187359 | Sequence 187359, |
| 535 | 328.3 | 4.4 | 448 | 1 | US-09-925-065A-870085 | Sequence 870085, |
| 536 | 328.3 | 4.4 | 448 | 1 | US-09-925-065A-870085 | Sequence 870085, |
| c 537 | 327.7 | 4.4 | 440 | 1 | US-09-925-065A-187360 | Sequence 187360, |
| c 538 | 327.7 | 4.4 | 440 | 1 | US-09-925-065A-187360 | Sequence 187360, |
| 539 | 326.7 | 4.4 | 448 | 1 | US-09-925-065A-870086 | Sequence 870086, |
| 540 | 326.7 | 4.4 | 448 | 1 | US-09-925-065A-870086 | Sequence 870086, |
| c 541 | 324.7 | 4.3 | 409 | 1 | US-09-864-761-4153 | Sequence 4153, Ap |
| c 542 | 324.2 | 4.3 | 452 | 1 | US-10-301-480-393751 | Sequence 393751, |
| c 543 | 324.2 | 4.3 | 452 | 1 | US-10-301-480-1007160 | Sequence 1007160, |
| c 544 | 323.8 | 4.3 | 450 | 1 | US-09-925-065A-399985 | Sequence 399985, |

| | | | | | | |
|---|---|---|---|---|---|---|
| c 545 | 323.8 | 4.3 | 450 | 1 | US-09-925-065A-399985 | Sequence 399985, |
| c 546 | 322.2 | 4.3 | 450 | 1 | US-09-925-065A-318955 | Sequence 318955, |
| c 547 | 322.2 | 4.3 | 450 | 1 | US-09-925-065A-318955 | Sequence 318955, |
| 548 | 322 | 4.3 | 452 | 1 | US-09-925-065A-863188 | Sequence 863188, |
| 549 | 322 | 4.3 | 452 | 1 | US-09-925-065A-863188 | Sequence 863188, |
| c 550 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-890779 | Sequence 890779, |
| c 551 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-890779 | Sequence 890779, |
| c 552 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-914003 | Sequence 914003, |
| c 553 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-914003 | Sequence 914003, |
| c 554 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-914004 | Sequence 914004, |
| c 555 | 321.5 | 4.3 | 449 | 1 | US-09-925-065A-914004 | Sequence 914004, |
| 556 | 320.4 | 4.3 | 441 | 1 | US-09-925-065A-893796 | Sequence 893796, |
| 557 | 320.4 | 4.3 | 441 | 1 | US-09-925-065A-893796 | Sequence 893796, |
| 558 | 320.4 | 4.3 | 441 | 1 | US-09-925-065A-893797 | Sequence 893797, |
| 559 | 320.4 | 4.3 | 441 | 1 | US-09-925-065A-893797 | Sequence 893797, |
| 560 | 320.4 | 4.3 | 452 | 1 | US-09-925-065A-863189 | Sequence 863189, |
| 561 | 320.4 | 4.3 | 452 | 1 | US-09-925-065A-863189 | Sequence 863189, |
| c 562 | 319.9 | 4.3 | 449 | 1 | US-09-925-065A-890778 | Sequence 890778, |
| c 563 | 319.9 | 4.3 | 449 | 1 | US-09-925-065A-890778 | Sequence 890778, |
| c 564 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-879842 | Sequence 879842, |
| c 565 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-879842 | Sequence 879842, |
| c 566 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-907856 | Sequence 907856, |
| c 567 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-907856 | Sequence 907856, |
| c 568 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-907857 | Sequence 907857, |
| c 569 | 319.1 | 4.3 | 446 | 1 | US-09-925-065A-907857 | Sequence 907857, |
| 570 | 318.8 | 4.3 | 441 | 1 | US-09-925-065A-893798 | Sequence 893798, |
| 571 | 318.8 | 4.3 | 441 | 1 | US-09-925-065A-893798 | Sequence 893798, |
| c 572 | 318.8 | 4.3 | 445 | 1 | US-10-301-480-467328 | Sequence 467328, |
| c 573 | 318.8 | 4.3 | 445 | 1 | US-10-301-480-1080737 | Sequence 1080737, |
| c 574 | 315.5 | 4.2 | 414 | 1 | US-09-864-761-4227 | Sequence 4227, Ap |
| c 575 | 309.6 | 4.1 | 440 | 1 | US-09-925-065A-147128 | Sequence 147128, |
| c 576 | 309.6 | 4.1 | 440 | 1 | US-09-925-065A-147128 | Sequence 147128, |
| 577 | 306.3 | 4.1 | 374 | 1 | US-10-040-916-47 | Sequence 47, Appl |
| c 578 | 292.1 | 3.9 | 336 | 1 | US-09-864-761-24312 | Sequence 24312, A |
| c 579 | 285.7 | 3.8 | 387 | 1 | US-09-864-761-30194 | Sequence 30194, A |
| c 580 | 279.4 | 3.7 | 332 | 1 | US-09-925-065A-817633 | Sequence 817633, |
| c 581 | 279.4 | 3.7 | 332 | 1 | US-09-925-065A-817633 | Sequence 817633, |
| c 582 | 268.4 | 3.6 | 379 | 1 | US-09-864-761-10242 | Sequence 10242, A |
| 583 | 262.7 | 3.5 | 310 | 1 | US-09-864-761-18530 | Sequence 18530, A |
| c 584 | 256.8 | 3.4 | 309 | 1 | US-09-925-065A-829270 | Sequence 829270, |
| c 585 | 256.8 | 3.4 | 309 | 1 | US-09-925-065A-829270 | Sequence 829270, |
| 586 | 251.2 | 3.4 | 323 | 1 | US-10-450-763-20002 | Sequence 20002, A |
| 587 | 251 | 3.4 | 279 | 1 | US-09-925-065A-585177 | Sequence 585177, |
| 588 | 251 | 3.4 | 279 | 1 | US-09-925-065A-585177 | Sequence 585177, |
| 589 | 247.3 | 3.3 | 333 | 1 | US-10-029-386-17183 | Sequence 17183, A |
| 590 | 244.5 | 3.3 | 275 | 1 | US-09-864-761-30531 | Sequence 30531, A |
| c 591 | 243.1 | 3.2 | 317 | 1 | US-09-925-065A-114790 | Sequence 114790, |
| c 592 | 243.1 | 3.2 | 317 | 1 | US-09-925-065A-114790 | Sequence 114790, |
| 593 | 238 | 3.2 | 279 | 1 | US-10-040-916-50 | Sequence 50, Appl |
| 594 | 237.1 | 3.2 | 310 | 1 | US-08-979-847-100 | Sequence 100, App |
| 595 | 237.1 | 3.2 | 310 | 1 | US-10-114-104-100 | Sequence 100, App |
| c 596 | 236 | 3.2 | 266 | 1 | US-10-029-386-14268 | Sequence 14268, A |
| 597 | 235.6 | 3.1 | 290 | 1 | US-10-242-535A-31708 | Sequence 31708, A |
| 598 | 235.6 | 3.1 | 290 | 1 | US-10-085-783A-31708 | Sequence 31708, A |
| 599 | 215.2 | 2.9 | 297 | 1 | US-08-979-847-2 | Sequence 2, Appli |
| 600 | 215.2 | 2.9 | 297 | 1 | US-10-430-442-2 | Sequence 2, Appli |
| 601 | 215.2 | 2.9 | 297 | 1 | US-10-114-104-2 | Sequence 2, Appli |
| 602 | 205 | 2.7 | 278 | 1 | US-09-864-761-17758 | Sequence 17758, A |
| 603 | 181.8 | 2.4 | 201 | 1 | US-10-741-600-65917 | Sequence 65917, A |
| c 604 | 177.2 | 2.4 | 200 | 1 | US-09-864-761-20979 | Sequence 20979, A |
| c 605 | 172.2 | 2.3 | 201 | 1 | US-11-124-368A-7793 | Sequence 7793, Ap |
| c 606 | 171 | 2.3 | 201 | 1 | US-11-124-368A-7851 | Sequence 7851, Ap |
| c 607 | 164 | 2.2 | 201 | 1 | US-10-741-600-46190 | Sequence 46190, A |
| c 608 | 164 | 2.2 | 201 | 1 | US-10-995-561-50894 | Sequence 50894, A |
| c 609 | 161.9 | 2.2 | 201 | 1 | US-10-995-561-50897 | Sequence 50897, A |
| c 610 | 161.3 | 2.2 | 201 | 1 | US-10-995-561-50899 | Sequence 50899, A |
| 611 | 156.6 | 2.1 | 202 | 1 | US-09-864-761-32986 | Sequence 32986, A |
| 612 | 156.2 | 2.1 | 207 | 1 | US-09-864-761-30401 | Sequence 30401, A |
| c 613 | 154.7 | 2.1 | 201 | 1 | US-11-124-368A-7852 | Sequence 7852, Ap |
| c 614 | 154.3 | 2.1 | 201 | 1 | US-10-741-600-46337 | Sequence 46337, A |
| c 615 | 154.3 | 2.1 | 201 | 1 | US-10-995-561-51113 | Sequence 51113, A |
| c 616 | 150.7 | 2.0 | 201 | 1 | US-10-741-600-46338 | Sequence 46338, A |
| c 617 | 150.7 | 2.0 | 201 | 1 | US-10-995-561-51115 | Sequence 51115, A |
| 618 | 147.5 | 2.0 | 201 | 1 | US-10-719-993-20008 | Sequence 20008, A |
| 619 | 147.5 | 2.0 | 201 | 1 | US-10-719-993-53980 | Sequence 53980, A |
| 620 | 147.5 | 2.0 | 201 | 1 | US-10-741-600-39842 | Sequence 39842, A |
| 621 | 147.5 | 2.0 | 201 | 1 | US-10-741-600-71390 | Sequence 71390, A |
| c 622 | 145.2 | 1.9 | 169 | 1 | US-10-029-386-18012 | Sequence 18012, A |
| c 623 | 141.9 | 1.9 | 194 | 1 | US-10-029-386-25984 | Sequence 25984, A |
| 624 | 139.2 | 1.9 | 150 | 1 | US-08-979-847-36 | Sequence 36, Appl |
| 625 | 139.2 | 1.9 | 150 | 1 | US-10-114-104-36 | Sequence 36, Appl |
| 626 | 135.6 | 1.8 | 174 | 1 | US-10-029-386-17995 | Sequence 17995, A |
| 627 | 122.6 | 1.6 | 152 | 1 | US-09-864-761-30590 | Sequence 30590, A |
| 628 | 121.8 | 1.6 | 133 | 1 | US-09-864-761-33021 | Sequence 33021, A |
| 629 | 121 | 1.6 | 155 | 1 | US-09-864-761-18806 | Sequence 18806, A |
| c 630 | 101.8 | 1.4 | 135 | 1 | US-09-864-761-26323 | Sequence 26323, A |
| 631 | 98 | 1.3 | 130 | 1 | US-10-310-914A-9963 | Sequence 9963, Ap |
| c 632 | 93 | 1.2 | 110 | 1 | US-09-864-761-23758 | Sequence 23758, A |
| c 633 | 89.2 | 1.2 | 103 | 1 | US-09-864-761-26878 | Sequence 26878, A |
| 634 | 85.2 | 1.1 | 107 | 1 | US-09-864-761-17757 | Sequence 17757, A |
| c 635 | 85 | 1.1 | 104 | 1 | US-10-029-386-19320 | Sequence 19320, A |
| 636 | 79.2 | 1.1 | 111 | 1 | US-08-979-847-7 | Sequence 7, Appli |
| 637 | 79.2 | 1.1 | 111 | 1 | US-10-430-442-7 | Sequence 7, Appli |
| 638 | 79.2 | 1.1 | 111 | 1 | US-10-114-104-7 | Sequence 7, Appli |

ALIGNMENTS

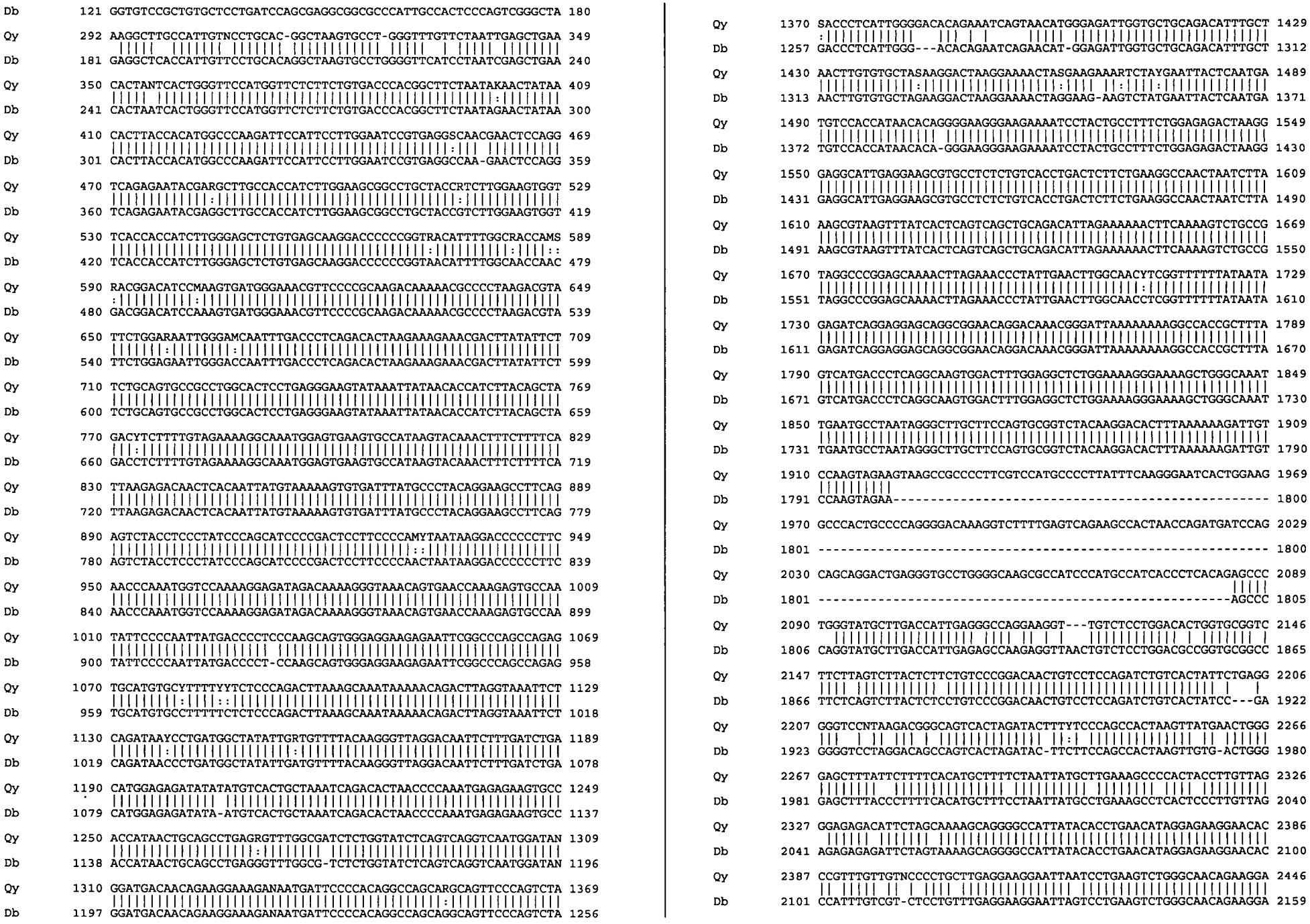
```
RESULT 1
US-10-450-763-12279
; Sequence 12279, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 12279
;   LENGTH: 4349
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (466)..(906)
;   OTHER INFORMATION: 100% homologous to Homo sapiens neurotensin
;   OTHER INFORMATION: receptor,accession number X70070,Smith-Waterman Score=775.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(4349)
;   OTHER INFORMATION: n = a,t,c or g
US-10-450-763-12279

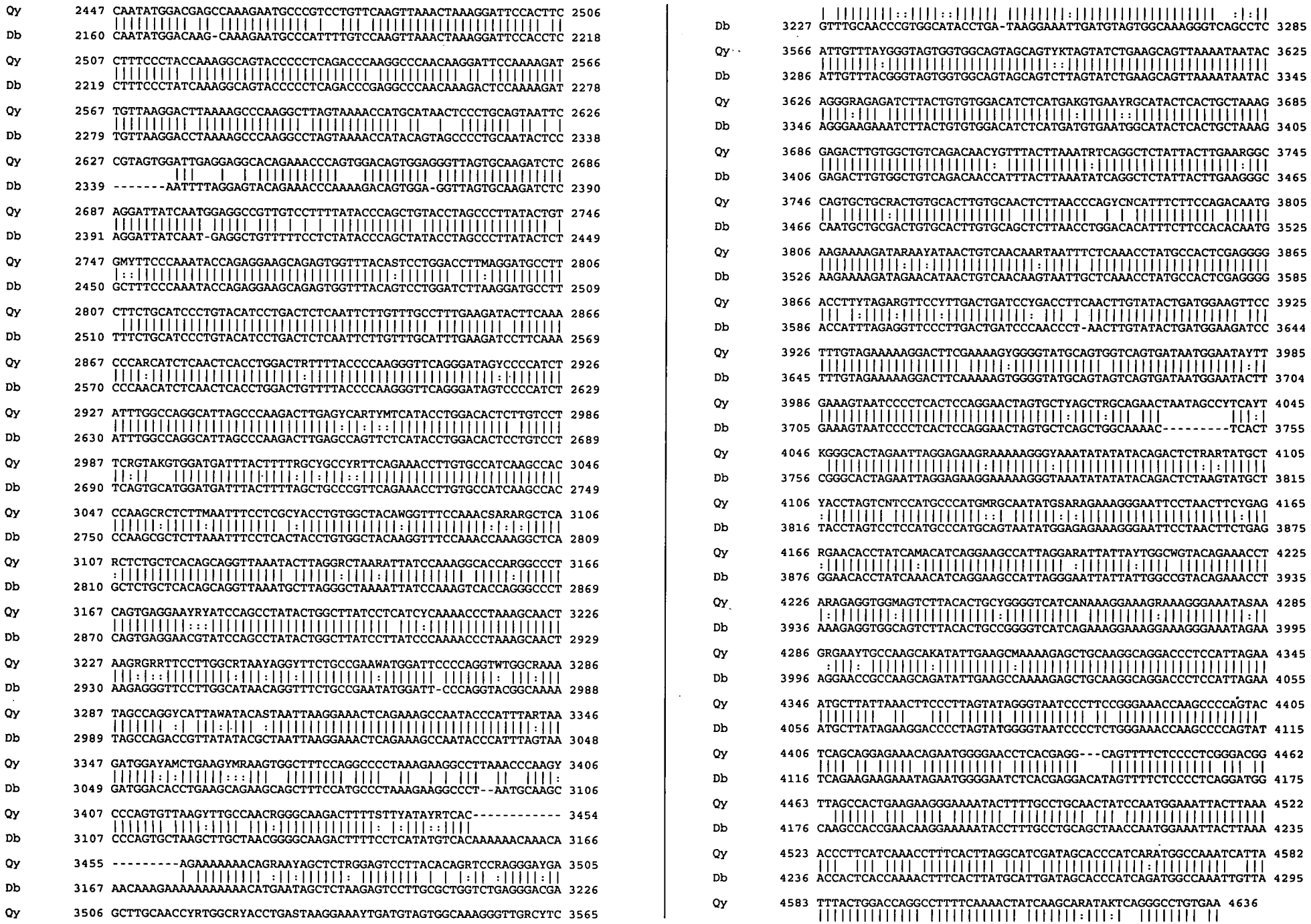
  Query Match             47.4%;  Score 3549.5;  DB 1;  Length 4349;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches 3974;  Conservative  148;  Mismatches  195;  Indels  237;  Gaps   28;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGA---RGCGCCCATTGCCGCTCCCAATTGGGCTA 291
              |||||||||||||||||||||||||||   :|||||||||||| |||||| | ||||||
```





```
Db       4296 TTTACTGGACCAGGACTTTTCAAAACTCTCAAGCAGATAGTTGGGGCCTGTAAA 4349


RESULT 2
US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
;  APPLICANT: Keith, Jr., James C.
;  APPLICANT:  McCoy, John M.
;  APPLICANT:  Mi, Sha
;  TITLE OF INVENTION: Methods and compositions for diagnosing
;  TITLE OF INVENTION:  and treating preeclampsia and gestational trophoblast
;  TITLE OF INVENTION:  disorders
;  FILE REFERENCE: GIN-6006B4
;  CURRENT APPLICATION NUMBER: US/09/902,535
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: 60/216,657
;  PRIOR FILING DATE: 2000-07-06
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (930)...(2546)
US-09-902-535-1

  Query Match             38.6%;  Score 2890.1;  DB 1;  Length 2930;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2873;  Conservative   48;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:|||||||||||||||||:|:||:||
Db         ,8 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 67

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 127

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        128 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 187

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 247

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 307

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 367

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 427

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 487

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 547

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 607

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             ||||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||
Db        608 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 667

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 727

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 787

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 847

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        848 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 907

Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
             |||||||||||||||| ||||||||:|||||||||||||||||||||:||:||||:
Db        908 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 966

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1026

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1086

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1146

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1206

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1266

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1326

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1386

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1446

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1506

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1566

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1567 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1626

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1686
```

```
Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1746

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1806

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
             |||||||||:|||:|||||||||||||||||||||||||| ||||||||||||||||||
Db       1807 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1866

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1926

Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1986

Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2046

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
             ||||||||||||||||||||||:||||||||||||||||||:|||||||||||||||:|
Db       2047 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2106

Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2166

Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
             |||||||||||||||:|||||||||:||||||||||||||||:||||||:|||||||
Db       2167 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2226

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2227 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2286

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2287 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2346

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2347 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2406

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       2407 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2466

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2526

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2527 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2586

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2587 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2646

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
             ||:|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       2647 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2706

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2707 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2766

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
```

```
             ||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
Db       2767 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2826

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2886

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||||||||||||||||||||||||||||:||||
Db       2887 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 2930


RESULT 3
US-10-114-893-134
; Sequence 134, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT: McCoy, John M.
;  APPLICANT: LaVallie, Edward R.
;  APPLICANT: Collins-Racie, Lisa A.
;  APPLICANT: Evans, Cheryl
;  APPLICANT: Merberg, David
;  APPLICANT: Treacy, Maurice
;  APPLICANT: Bowman, Michael R.
;  APPLICANT: Spaulding, Vikki
;  APPLICANT: Carlin-Duckett, McKeough
;  APPLICANT: Kelleher, Kerry S.
;  APPLICANT: Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: GI 6000-10A
;  CURRENT APPLICATION NUMBER: US/10/114,893
;  CURRENT FILING DATE: 2002-04-02
;  EARLIER APPLICATION NUMBER: 09/413,232
;  EARLIER FILING DATE: 1999-10-06
;  NUMBER OF SEQ ID NOS: 321
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
;   LENGTH: 2946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-114-893-134

  Query Match            38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             |||||||:|||||||||||||||||||:|||||||:|||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

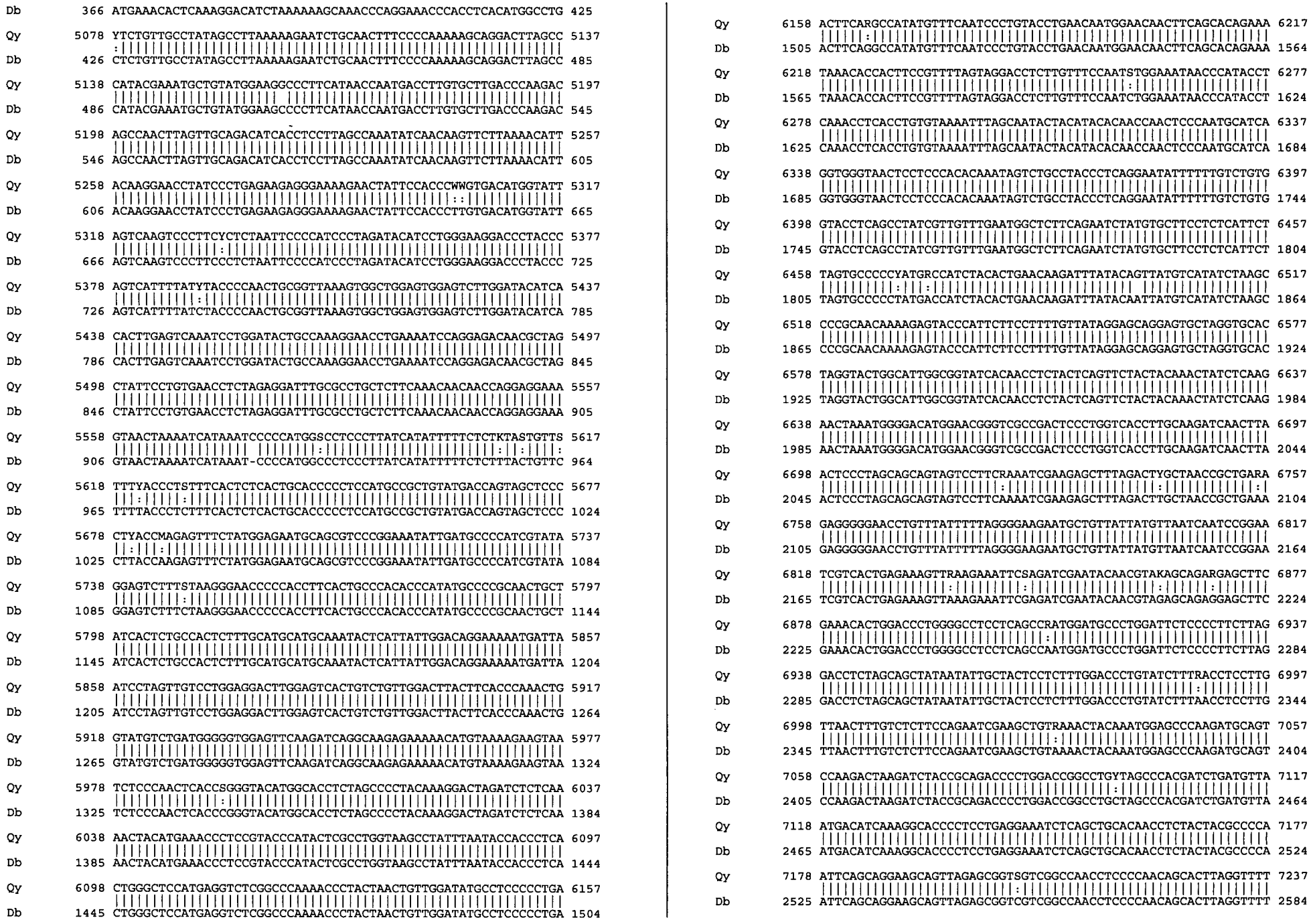
Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



```
Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
             ||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
             ||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db       2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||||||||||||||||:||||:
Db       2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929


RESULT 4
US-10-016-249-3
; Sequence 3, Application US/10016249
; Publication No. US20030100053A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT:  McCoy, John M.
;  APPLICANT:  LaVallie, Edward R.
;  APPLICANT:  Collins-Racie, Lisa A.
;  APPLICANT:  Evans, Cheryl
;  APPLICANT:  Merberg, David
;  APPLICANT:  Treacy, Maurice
;  APPLICANT:  Mi, Sha
;  APPLICANT:  Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: 6006B.AJ172A
;  CURRENT APPLICATION NUMBER: US/10/016,249
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: US/09/175,928
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-016-249-3

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             ||||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||
Db         606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db         666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905

Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
             |||||||||||||||||| |||||||:|||||||||||||||||||||:||:||||:
Db         906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
```

```
Db       1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097

Db       1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157

Db       1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217

Db       1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277

Db       1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337

Db       1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684

Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397

Db       1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457

Db       1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517

Db       1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577

Db       1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924

Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637

Db       1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984

Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697

Db       1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757

Db       2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104

Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817

Db       2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164

Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877

Db       2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937

Db       2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997

Db       2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057

Db       2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117

Db       2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177

Db       2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237

Db       2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297

Db       2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357

Db       2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417

Db       2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477

Db       2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537

Db       2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582

Db       2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929
```

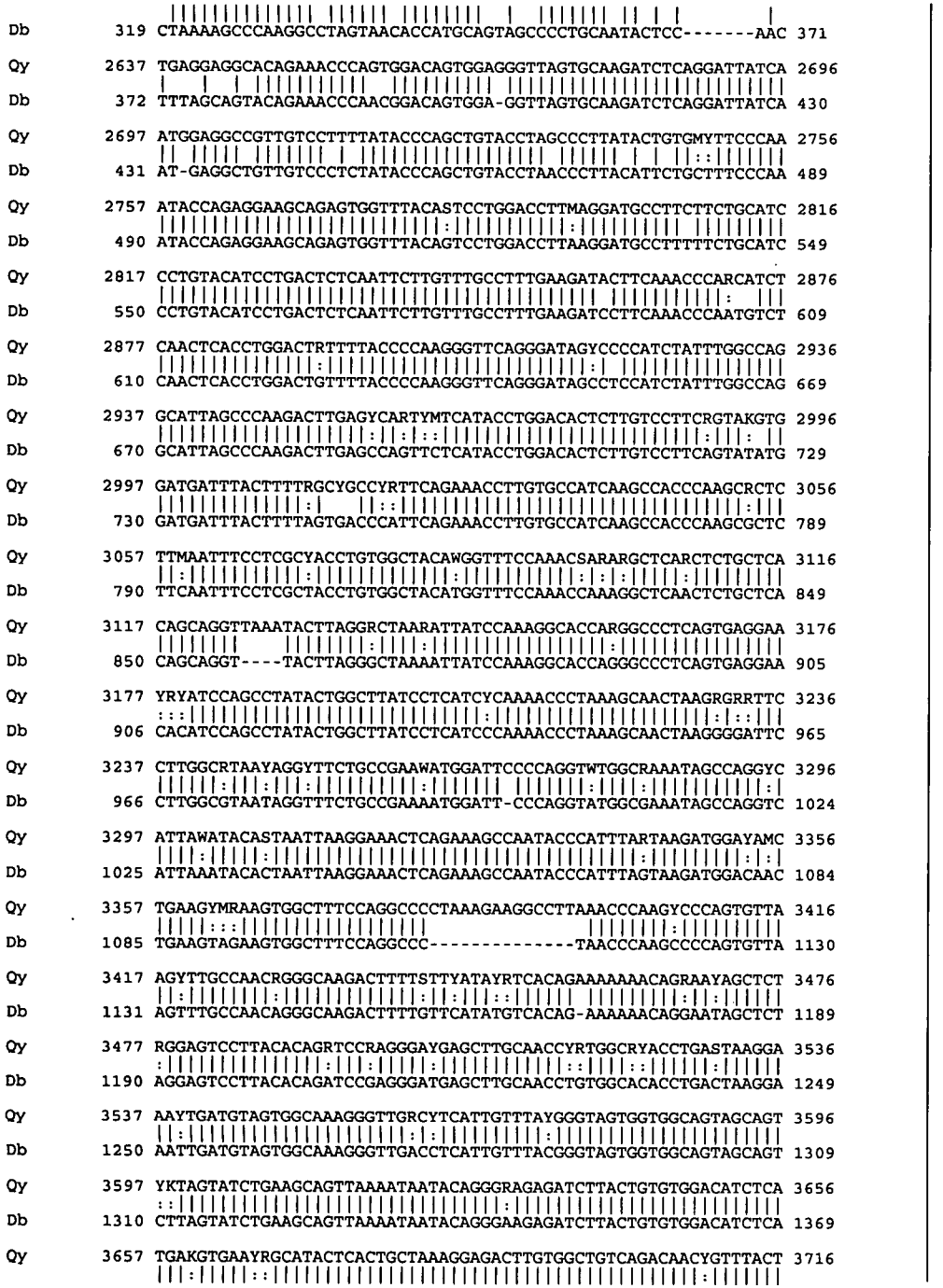
```
RESULT 5
US-10-450-763-7531
; Sequence 7531, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 7531
;   LENGTH: 3831
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (262)..(858)
;   OTHER INFORMATION: 85% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=895.
US-10-450-763-7531

  Query Match             36.7%;  Score 2750.9;  DB 1;  Length 3831;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches 2814;  Conservative  140;  Mismatches   54;  Indels   33;  Gaps   10;

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576

Db        259 CATCGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCAAAAAAATTGTTAAGGAC 318

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
```


Db 319 CTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCAATACTCC-------AAC 371
Qy 2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCA 2696
Db 372 TTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCA 430
Qy 2697 ATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAA 2756
Db 431 AT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTACATTCTGCTTTCCCAA 489
Qy 2757 ATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATC 2816
Db 490 ATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATC 549
Qy 2817 CCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCT 2876
Db 550 CCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCT 609
Qy 2877 CAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAG 2936
Db 610 CAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCTCCATCTATTTGGCCAG 669
Qy 2937 GCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTG 2996
Db 670 GCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCAGTATATG 729
Qy 2997 GATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTC 3056
Db 730 GATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTC 789
Qy 3057 TTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCA 3116
Db 790 TTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCA 849
Qy 3117 CAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAA 3176
Db 850 CAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAA 905
Qy 3177 YRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTC 3236
Db 906 CACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGGGGATTC 965
Qy 3237 CTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYC 3296
Db 966 CTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAATAGCCAGGTC 1024
Qy 3297 ATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMC 3356
Db 1025 ATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACAAC 1084
Qy 3357 TGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTA 3416
Db 1085 TGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGCCCCAGTGTTA 1130
Qy 3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
Db 1131 AGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAGGAATAGCTCT 1189
Qy 3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
Db 1190 AGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACCTGACTAAGGA 1249
Qy 3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
Db 1250 AATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGT 1309
Qy 3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
Db 1310 CTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCA 1369
Qy 3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
Db 1370 TGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTACT 1429
Qy 3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
Db 1430 TAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAACTCT 1489
Qy 3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
Db 1490 TAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACTGTCAACAAGT 1549
Qy 3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
Db 1550 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGACTGATCCCGA 1609
Qy 3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
Db 1610 CC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGTGGGGT 1668
Qy 3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db 1669 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 1728
Qy 4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db 1729 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 1788
Qy 4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db 1789 CAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 1847
Qy 4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db 1848 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 1907
Qy 4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db 1908 GGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACACTGCCGGGGTC 1967
Qy 4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
Db 1968 ATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATTGAAGCCAAAA 2027
Qy 4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db 2028 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 2087
Qy 4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db 2088 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 2147
Qy 4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
Db 2148 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 2207
Qy 4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
Db 2208 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 2267
Qy 4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
Db 2268 GCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 2327
Qy 4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
Db 2328 AGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTATCGCCAAGCTC 2387
Qy 4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
Db 2388 CTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTTTACCCA 2447
Qy 4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
Db 2448 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTTTCACGGGTTG 2507

```
Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2508 GGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 2567

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2568 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 2627

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2628 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 2687

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       2688 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGACA 2747

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db       2748 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGCCTATAGCCTT 2807

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2808 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 2867

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2868 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 2927

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2928 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 2987

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
             |||||||  |||  | ||||||||||::|||||||||||||||||||||||||||:||||
Db       2988 GAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTCCCTTCCCTCT 3045

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       3046 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCA 3105

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3106 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 3165

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3166 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 3225

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             |||||||||||||||||||||||||||||||||||||||||
Db       3226 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 3266


RESULT 6
US-10-632-793-21
; Sequence 21, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
;   LENGTH: 2938
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-21

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              |||||||||||||:||||::||||||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
              ||||||||||:|||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy       1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

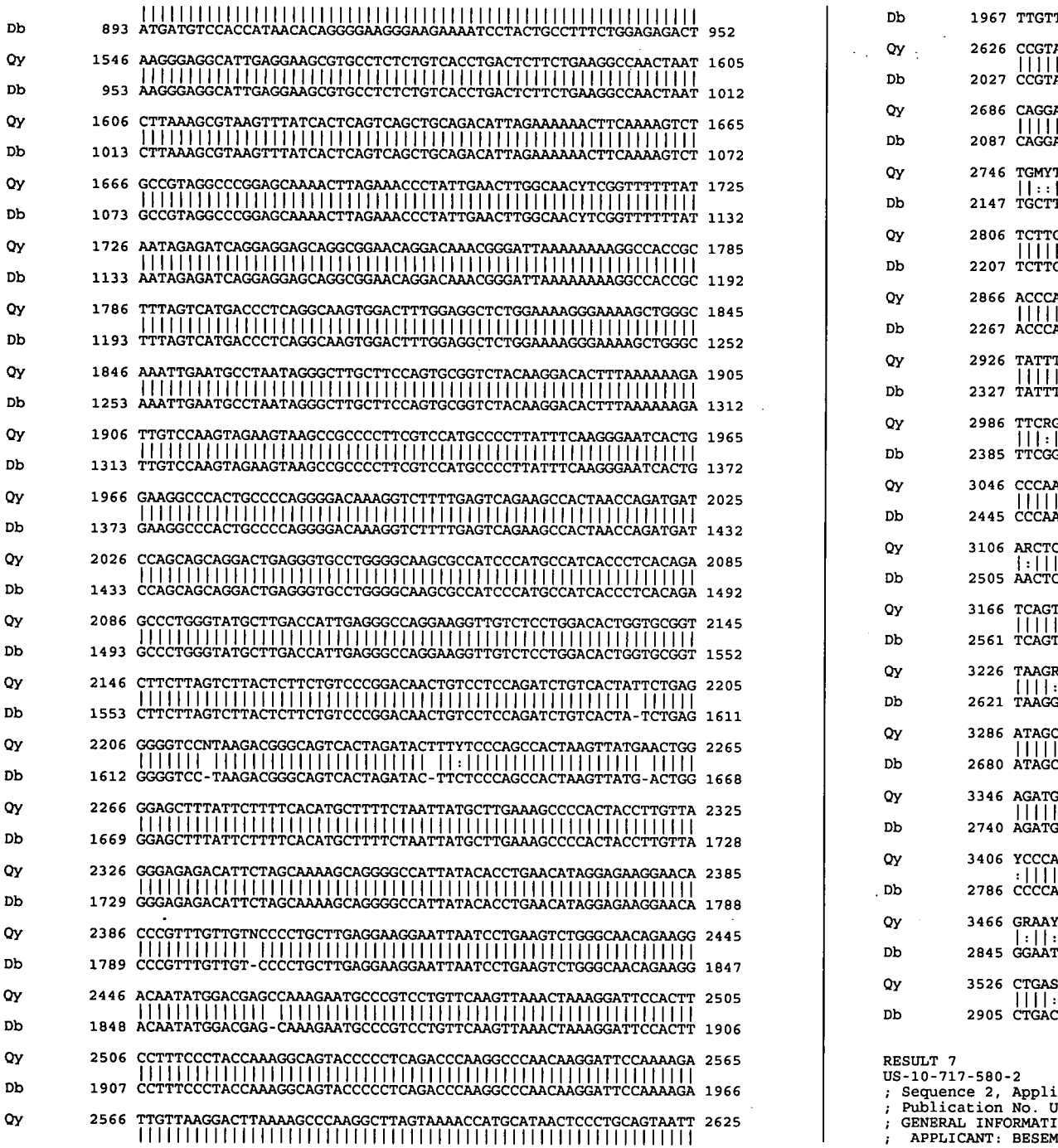
Qy       1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy       1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
                 |||||||||||||| | ||  |  |||  |||||||:||||||||||||||||||||
Db        718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy       1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy       1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
              ||||||||||||||||:||||||||||||||||||:|||||||:||||:||||||||||||
Db        833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy       1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
```




RESULT 7
US-10-717-580-2
; Sequence 2, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic

```
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2938
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-2

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              |||||||||||||:||||::|||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
              ||||||||||:||||||||||||||:||||||||||||||||||||||||||||||||||
Db        540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy       1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

Qy       1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
              |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy       1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
              |||||||||||||| | ||    |       |||    |||||||:||||||||||
Db        718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy       1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy       1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
              ||||||||||||||||:||||||||||||||||:|||||||:||||:||||||||||||
Db        833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy       1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952

Qy       1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012

Qy       1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072

Qy       1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132

Qy       1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192

Qy       1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252

Qy       1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552

Qy       2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611

Qy       2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
              ||||||| |||||||||||||||||||||||| ||:|||||||||||||||||| |||||
Db       1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668

Qy       2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728

Qy       2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788

Qy       2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
             |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847

Qy       2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906

Qy       2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966

Qy       2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026

Qy       2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086

Qy       2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db       2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||||:|||||||
Db       2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             |||||||||||||||||||||||||||||||:||:|::|||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:||||||||||:||||||||||||:||||||||||:|:|:|||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||||||    ||||||||:||||:|||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:||||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||||||||||||||||||||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :||||||||||||:|||||||:||||||||||||||:||:||||::|||||| |||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:||||||:|||||||||||||||||:|||:||||||:||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 8
US-10-632-793-28
; Sequence 28, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
;   LENGTH: 3372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-28

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:|||||||||||||||||||||||||||:|||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

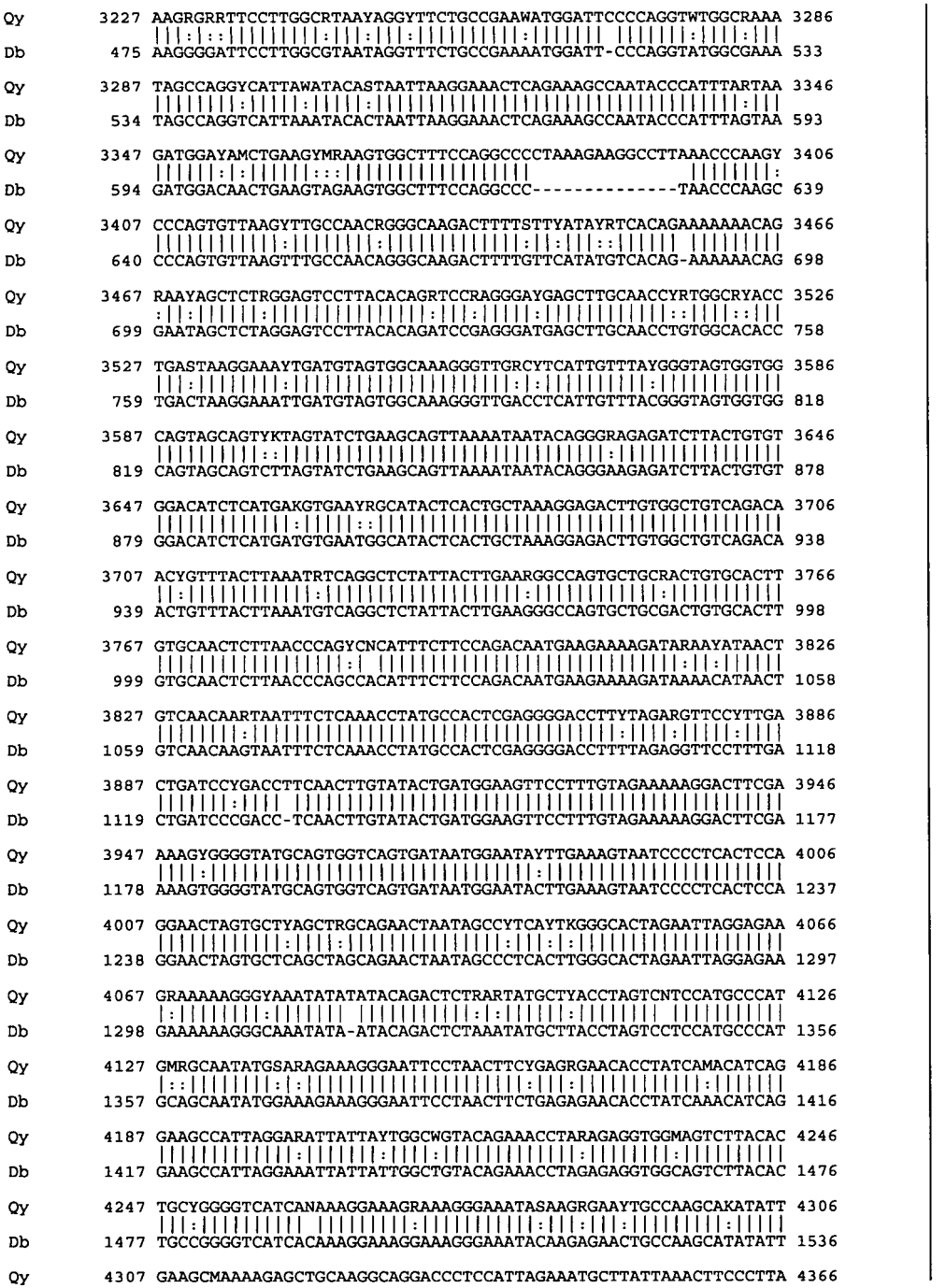
Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||||:||:|::|||||||||||||  ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:||||||||||:||||||||||||:||||||||||:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||    ||||||||:||||:|||||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474
```



```
Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785


RESULT 9
US-10-717-580-9
; Sequence 9, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 3372
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-9

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:|||||||||||||||||||||:|||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||||:||:|::|||||||||||  ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||:||:|||::|||||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:|||||||||||:|||||||||||||:|||||||||||:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||    ||||||||:||||:|||||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||:||||||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::|||||||||:|||:|||:||||||||||:|||||||  |||||||:||||:|||
Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db        534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
             ||||||:|:||||||:::|||||||||||||||||||||||||||||||||||||||||:
Db        594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC----------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
             ||||||||||:||||||||:|||||||||||||:||:|||::||||||  |||||||||
Db        640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
             :||:|||||||:|||||||||||||||:|||:|||||:||||||||||||::||||::|||
Db        699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
             |||:||||||||:|||||||||||||||||||||:|||||||||||:||||||||||||||
Db        759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
             ||||||||||::|||||||||||||||||||||||||||||||:||||||||||||||||
Db        819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
             |||||||||||||:|||||::||||||||||||||||||||||||||||||||||||||||
Db        879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
             ||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
             |||||||||||||||||:| ||||||||||||||||||||||||||:||:||||||||||
Db        999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||||:|||||||||||||||||||||||||||||||||||||:||||:|||||:||||
Db       1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
             ||||||||||||:||||:|||||||||||||||:|||:|:|||||||||||||||||||||
Db       1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
             |:|||||||:|||||| ||||||||||:|:|||||:|||||||| |||||||||||||||
Db       1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356

Qy       4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
             |::|||||||:|:|||||||||||||||||||||||||||||:|||:|||||||||||||
Db       1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416

Qy       4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
             ||||||||||||:|||||||:||||:|||||||||||||:|||||||||:||||||||||
Db       1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476

Qy       4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
```

```
Db       1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536

Qy       4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366

Db       1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596

Qy       4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426

Db       1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656

Qy       4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486

Db       1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716

Qy       4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546

Db       1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606

Db       1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666

Db       1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726

Db       1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956

Qy       4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786

Db       1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016

Qy       4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846

Db       2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076

Qy       4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906

Db       2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136

Qy       4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966

Db       2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196

Qy       4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026

Db       2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256

Qy       5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086

Db       2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316

Qy       5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146

Db       2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376

Qy       5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206

Db       2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266

Db       2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496

Qy       5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326

Db       2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554

Qy       5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386

Db       2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446

Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506

Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557

Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785
```

RESULT 10
US-11-028-539-1
; Sequence 1, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUM
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-028-539-1

Query Match 34.1%; Score 2551.4; DB 1; Length 2599;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2552; Conservative 39; Mismatches 5; Indels 2; Gaps 2;

```
Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714

Db          4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy       4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774

Db         64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123

Qy       4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834

Db        124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183

Qy       4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894

Db        184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243

Qy       4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954

Db        244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303

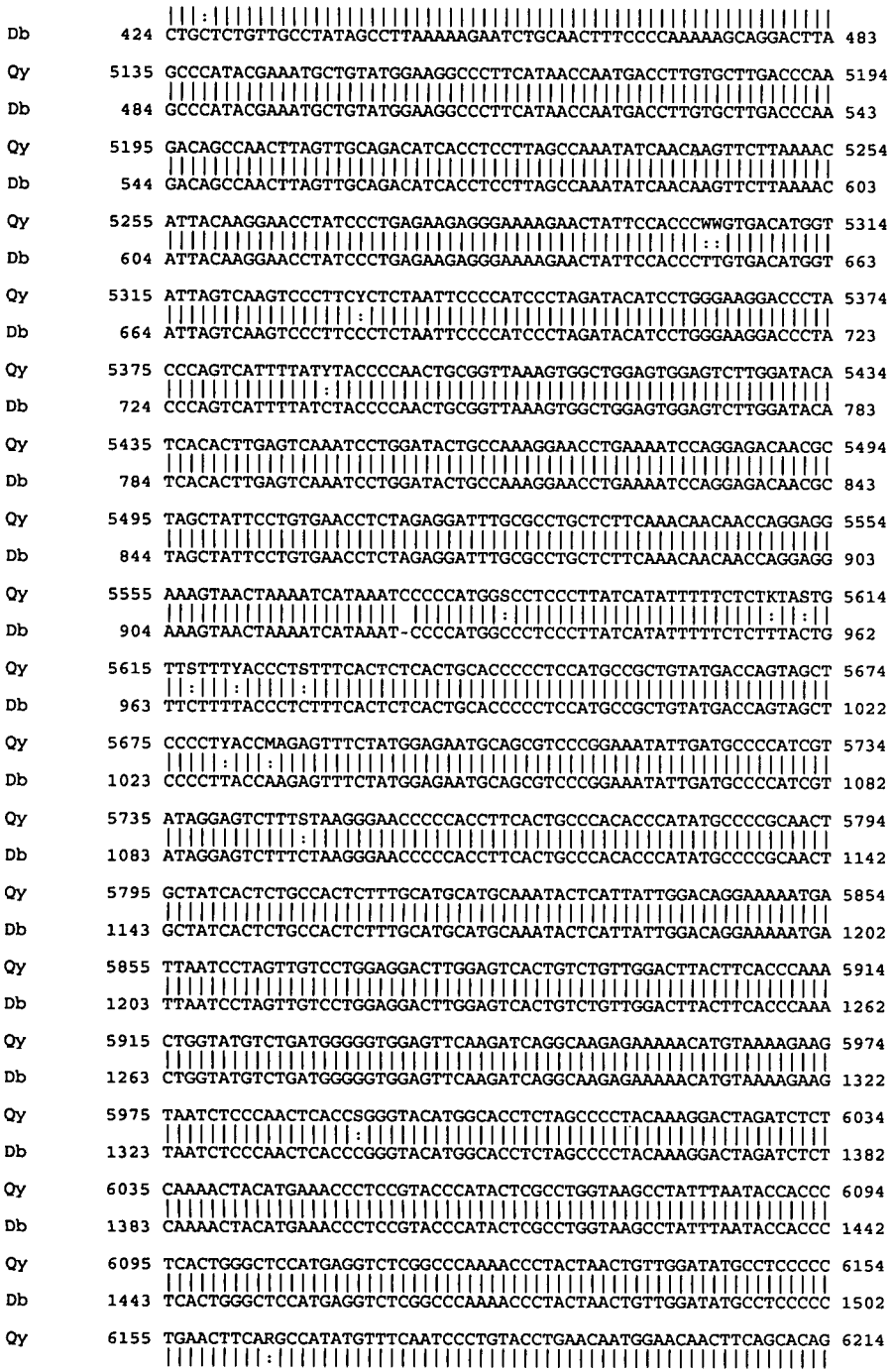
Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014

Db        304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363

Qy       5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074

Db        364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
```


Db 424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483
Qy 5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
Db 484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543
Qy 5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
Db 544 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 603
Qy 5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
Db 604 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 663
Qy 5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
Db 664 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 723
Qy 5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
Db 724 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 783
Qy 5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
Db 784 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 843
Qy 5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554
Db 844 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 903
Qy 5555 AAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTG 5614
Db 904 AAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTG 962
Qy 5615 TTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 5674
Db 963 TTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 1022
Qy 5675 CCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 5734
Db 1023 CCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 1082
Qy 5735 ATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 5794
Db 1083 ATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 1142
Qy 5795 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 5854
Db 1143 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 1202
Qy 5855 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 5914
Db 1203 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 1262
Qy 5915 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 5974
Db 1263 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 1322
Qy 5975 TAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 6034
Db 1323 TAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 1382
Qy 6035 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 6094
Db 1383 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 1442
Qy 6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 6154
Db 1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 1502
Qy 6155 TGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 6214
Db 1503 TGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 1562
Qy 6215 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATA 6274
Db 1563 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATA 1622
Qy 6275 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 6334
Db 1623 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 1682
Qy 6335 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 6394
Db 1683 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 1742
Qy 6395 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 6454
Db 1743 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 1802
Qy 6455 TCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 6514
Db 1803 TCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 1862
Qy 6515 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 6574
Db 1863 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 1922
Qy 6575 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 6634
Db 1923 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 1982
Qy 6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694
Db 1983 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 2042
Qy 6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754
Db 2043 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 2102
Qy 6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814
Db 2103 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 2162
Qy 6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874
Db 2163 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 2222
Qy 6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934
Db 2223 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 2282
Qy 6935 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCC 6994
Db 2283 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCC 2342
Qy 6995 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGC 7054
Db 2343 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGC 2402
Qy 7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATG 7114
Db 2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATG 2462
Qy 7115 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 7174
Db 2463 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 2522
Qy 7175 CCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGT 7234
Db 2523 CCAATTCAGCAGGAAGCAGTTAGAGCGGTC-TCGGCCAACCTCCCCAACAGCACTTAGGT 2581
Qy 7235 TTTCCTGTTGAGATGGGG 7252
Db 2582 TTTCCTGTTGAGATGGGG 2599

```
RESULT 11
US-11-028-539-4
; Sequence 4, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/11/028,539
;  CURRENT FILING DATE:  2005-01-05
;  PRIOR APPLICATION NUMBER: US/09/719,554
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 2784
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-028-539-4

  Query Match            30.9%;  Score 2315.4;  DB 1;  Length 2784;
  Best Local Similarity  89.9%;  Pred. No. 0;
  Matches 2523;  Conservative   39;  Mismatches  180;  Indels   66;  Gaps    9;

Qy       4674 CTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTA 4732
             |||||||||||:||||||:||||||||| |||||  :|:||:|||||||||| |||||||
Db          1 CTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAAGAGAAGACTGGCAACTAGATTTTA 60

Qy       4733 CCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGG 4792
             |||| | ||||||| |||||||||||||||||||||||| ||||||:|||||||||| |
Db         61 CCCATATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGATACTTTCACTG 120

Qy       4793 GTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAG 4852
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||      |
Db        121 GTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAA------CG 174

Qy       4853 TTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTG 4912
             |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||
Db        175 TTCATGAAATAATTCCCAGATTCGGACTTCCCCAAGGCTTACAGAGTGACAATGGCCCTG 234

Qy       4913 CTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTAC 4972
             ||||| |||| |||||||||| |||||||||||| |||||||||| |||||||| ||
Db        235 CTTTCAAGGCTACAGTAACCCAAGGAGTATCCCAGGTGTTAGGTATACAATATCACTCAC 294

Qy       4973 ACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAG 5032
             |||||||||| |||||||||||||||| ||||| |||||||||||  |||:|||||||
Db        295 ACTGCGCCTGGAGGCCACAGTCCTCAGGAAAGGTGGAGAAAATGAACAAAACACTCAAAT 354

Qy       5033 GACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAG 5092
             |||||||||||||| || |||||||||||||| |||||||||:|||||||||||||||
Db        355 GACATCTAAAAAAGCTAATCCAGGAAACCCACCTCGCATGGCCTGCTCTGTTGCCTATAG 414

Qy       5093 CCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152
             |||||  ||||||| | |||| |||||||||||||||||||| |||||| |||||||||
Db        415 CCTTACTAAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGTCCATACAAAATGCTGT 474

Qy       5153 ATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTG 5211
             |||||  |||||||| |||||||||| |||| |||||||| | ||||||||||||||||||
Db        475 ATGGACGGCCCTTCCTAACCAATGAACTTGGGCTTGACCGAGAGACAGCCAACTTAGTTG 534

Qy       5212 CAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATC 5271
             ||||||||| |||||||||||||||||||||| |||||||||||||||| ||| ||| ||
Db        535 CAGACATCATCTCCTTAGCCAAATATCAACAGGTTCTTAAAACATTACAGGGAGCCTGTC 594

Qy       5272 CCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTC 5331
             ||  |||||||||||| ||||||||||||||: |||||||||||||||||||||||||||

Db        595 CCCAAGAAGAGGGAAAGGAACTATTCCACCCTGGTGACATGGTATTAGTCAAGTCCCTTC 654

Qy       5332 YCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTA 5391
             :|||||||||||||||||||||||||||||||||||||||||||  |||||||| ||||||||:||
Db        655 CCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGAAACTACCCAGCCATTTTATCTA 714

Qy       5392 CCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAT 5451
             |||  ||| || |||||||||||||||| ||||||||||||||||||||||  |||||||
Db        715 CCCTAACGGCAGTTAAAGTGGCTGGAGCGGAGTCTTGGATACATCACACTCAAGTCAAAC 774

Qy       5452 CCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAAC 5511
             |||||||||||||||||||||  |||||||| ||||||| |||||||||||||||||||||
Db        775 CCTGGATACTGCCAAAGGAACTCAAAAATCCATGAGACAATGCTAGCTATTCCTGTGAAC 834

Qy       5512 CTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCAT 5571
             |||||||||| |||||||||||||||||  |||||||| |||||||||||||||||||| |
Db        835 CTCTAGAGGATCTGCGCCTGCTCTTCAAATGACAACCAGGGGGAAAGTAACTAAAATCGT 894

Qy       5572 AAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTC 5631
             ||||||||  |||:|||||||||||||||||||||:||:||||:| |:|||| :||||
Db        895 AAATCCCC--TGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTCTTACCCCCTTTC 952

Qy       5632 ACTCTCACTGCACCCCCTCCATGCCGCTGTA------------------TGACCAGTAG 5672
             |||||||||||||||| |||||||| ||| |                  |||||||
Db        953 ACTCTCACTGCACCCCGTCCATGCCACTGCACCCCGTCCATGCCCCGTCCATGCCAGTAG 1012

Qy       5673 CTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATC 5732
             |||||||:| |:|||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTCCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATT 1072

Qy       5733 GTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAA 5792
             |||||||||| | |:|||||||||||||||||||||||||||||||||||||||||| |||
Db       1073 GTATAGGAGTTTATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAA 1132

Qy       5793 CTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAT 5852
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAC 1192

Qy       5853 GATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCA 5912
             |||||||| |||||||||||||||||||  |||||||||||||||||||| ||||| ||
Db       1193 GATTAATCCCAGTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCA 1239

Qy       5913 AACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGA 5972
             ||  |||||||||||||||||||||||||||||||||||||||| |||||||||  |||| ||
Db       1240 TACCAGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGA 1299

Qy       5973 AGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCT 6032
             |||||||||||||||| ||| |||||||| ||||| || ||||||||||||||||||||||
Db       1300 AGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCT 1359

Qy       6033 CTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCAC 6092
             |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db       1360 CTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCAC 1419

Qy       6093 CCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCC 6152
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1420 CCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCC 1479

Qy       6153 CCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCAC 6212
             |||| |||| |:|||||  |||||||||| ||||||||||||||||||||||||||||||
Db       1480 CCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCAC 1539

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             |||||||||||||||| |||||||||| ||||   ||||||||:|||||||||||||||
Db       1540 AGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCA 1596

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             |||||||||||||||||||||||||||||||||||||||||||||| | ||||| ||||||||||
Db       1597 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATG 1656

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1657 CATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 1716
```

```
Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||||||||
Db       1717 CTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTC 1776

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
             |||||||||| ||||:||| |||||||||||||||||||||||||||||| | |||||||| |
Db       1777 ATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACC 1836

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1837 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 1896

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
               | | ||| |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1897 CGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTC 1956

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             |||||||||||||| ||||||||| ||||||| ||  ||||||||||||||||||||||
Db       1957 TCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCA 2016

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGC 6752
             ||||||||||||||||||||||||||||||:||||||||||||||||||:|||||||||
Db       2017 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGC 2076

Qy       6753 TGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATC 6812
               ||:|| |||||||||| |||||||||| |||| ||||||||| ||||||||||||||||
Db       2077 GGAAAGCGGGGGAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATC 2136

Qy       6813 CGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGA 6872
             ||||||| |||| |||||||||:||||||||| || ||||||| ||||||:|||| |:||
Db       2137 CGGAATCATCACCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGA 2196

Qy       6873 GCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTT 6932
             ||| | ||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       2197 GCTGCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTT 2256

Qy       6933 CTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCT 6992
             |||||||||||||||||||||||||| |||||||||||||||||||||||||:||||
Db       2257 CTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCT 2316

Qy       6993 CCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAAT 7040
             ||||||||| ||||||| |||||||||||||| ||:|||||              |||||
Db       2317 CCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAAT 2376

Qy       7041 GGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAG 7100
             ||||||| |||||||||||||| || ||| |||||||| | ||||||||||||||||||:|||
Db       2377 GGAGCCCCAGATGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAG 2436

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160
             |||| | |||||||||||||||||||||||||||||||||| |||||||||||| |||||||
Db       2437 CCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACA 2496

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220
             |||||||||||||||||||||||||||||||||||||||||| ||| || ||||||||||||||
Db       2497 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC 2556

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280
             ||||||| || |||||||||||||||||| |||||||||||||||||| || || ||||||
Db       2557 AACAGCAGTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTC 2616

Qy       7281 CTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACA 7340
             |||| |  |:|||||||:|||| ||||:|||||||||||| | ||||||||||||||||
Db       2617 CTAGACCAACTAAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACA 2676

Qy       7341 CGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAAT 7391
             | |||||||||||||||:|||| ||  |||||||         ||||||||||||||||
Db       2677 CCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAAT 2736

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTG 7439
             ||||||||||||| ||||||| ||||||||||||||||||:| |||
Db       2737 GCTAATTAGGCAAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTG 2784
```

```
RESULT 12
US-10-632-793-29
; Sequence 29, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
;   LENGTH: 2372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1191)..(1191)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1213)..(1213)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2089)..(2089)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2274)..(2274)
;   OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-29

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

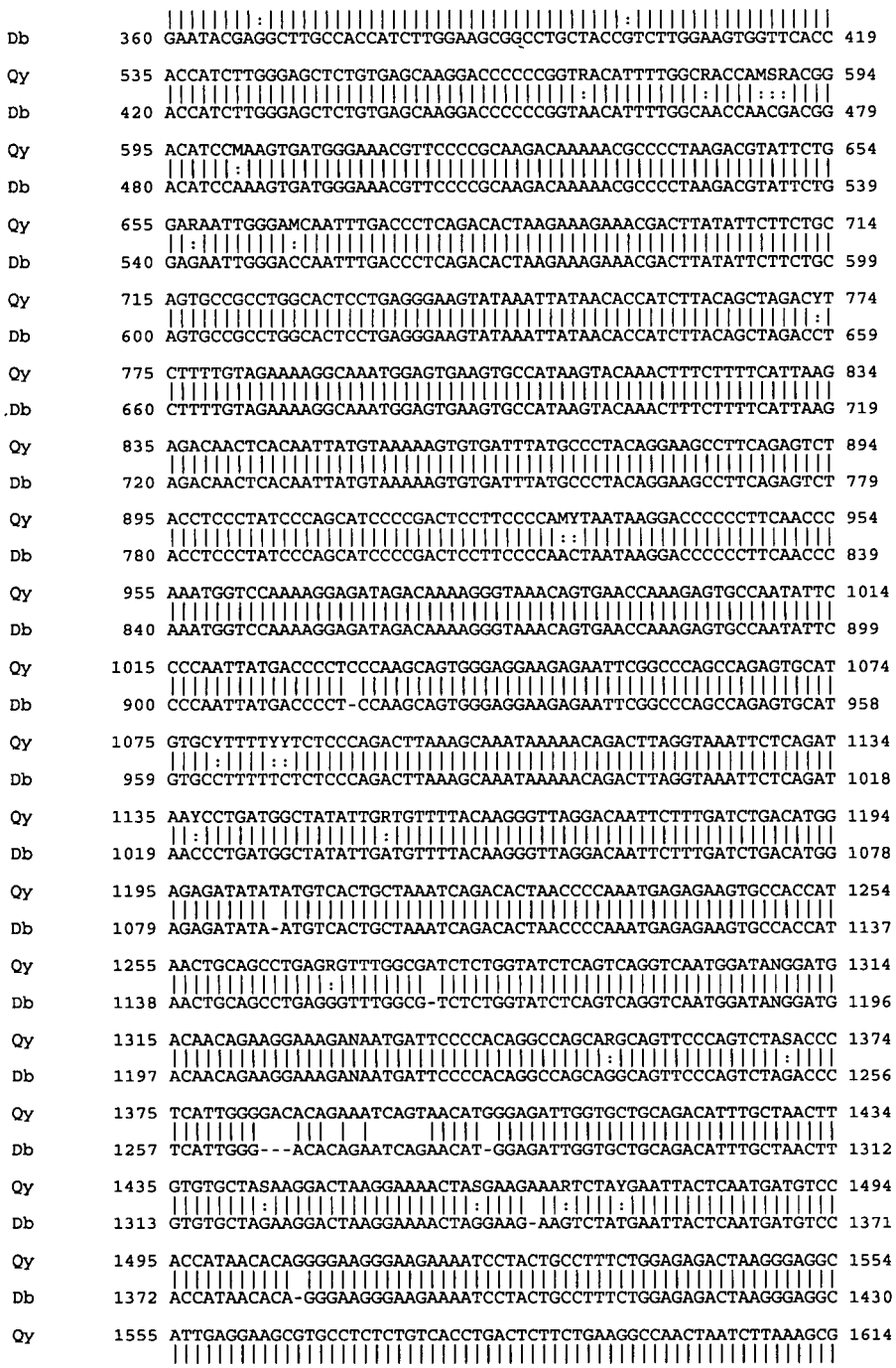
Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

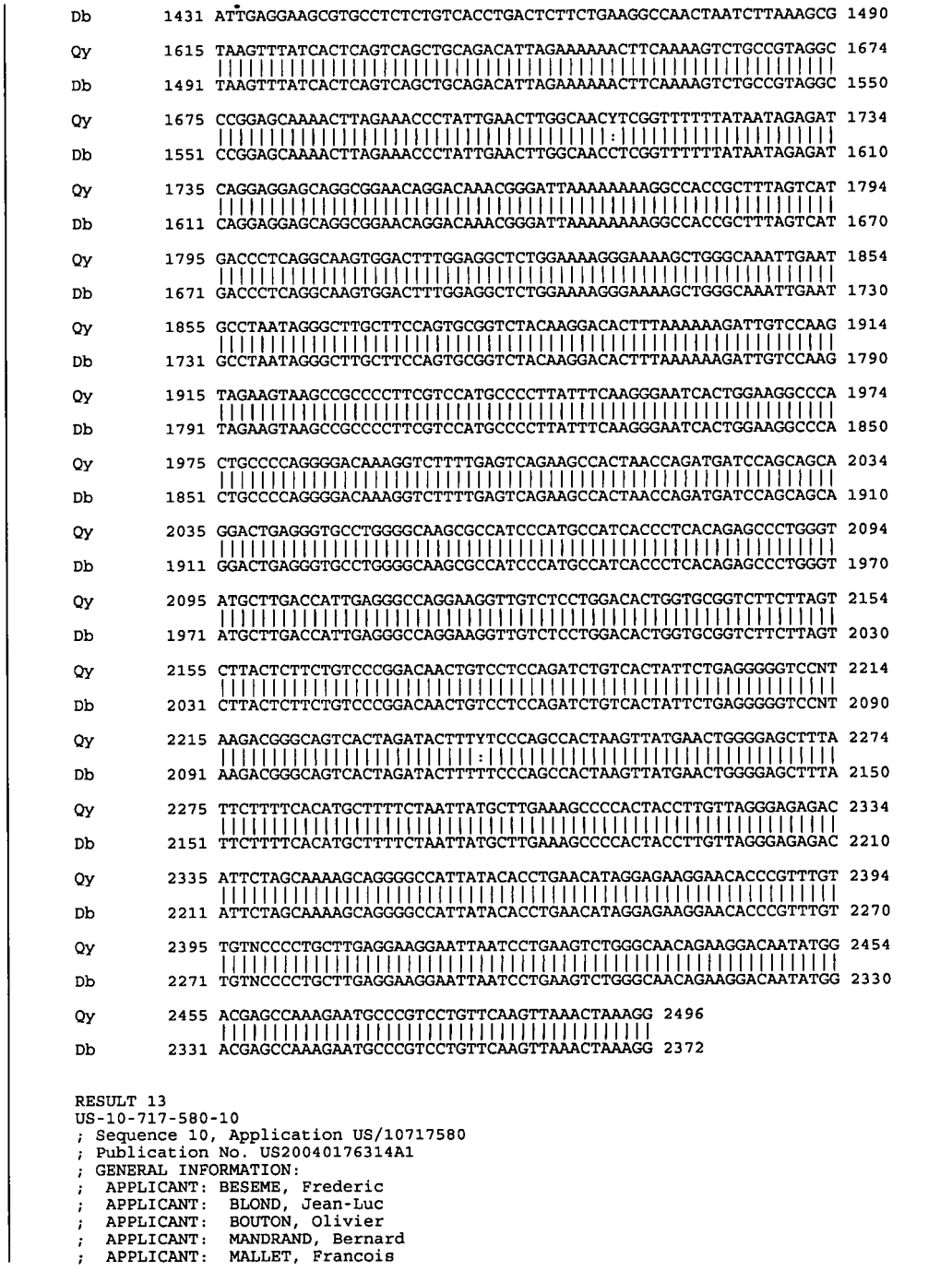
Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
```





RESULT 13
US-10-717-580-10
; Sequence 10, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois

```
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 2372
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1191)..(1191)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1213)..(1213)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2089)..(2089)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2274)..(2274)
;   OTHER INFORMATION: n = any nucleotide
US-10-717-580-10

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||::||||||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:|||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              |||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              ||||||||   ||| | |     |||||||||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:||||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
```

```
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
             |||||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372


RESULT 14
US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
;   LENGTH: 2782
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-26

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433
```

```
Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:|||||||:||||||||||||||||||||||||||||||||||:|||||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||:||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764
```

RESULT 15
US-10-717-580-7
; Sequence 7, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISF
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human
US-10-717-580-7

Query Match 29.4%; Score 2205.1; DB 1; Length 2782;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 2194; Conservative 36; Mismatches 0; Indels 1; Gaps 1;

```
Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833
```

```
Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764
```

RESULT 16
US-10-133-036-1
; Sequence 1, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659

```
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2782
;   TYPE: DNA
;   ORGANISM: Human endogenous retrovirus
US-10-133-036-1

  Query Match             29.4%;  Score 2201.9;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2192;  Conservative   36;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||:||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:|||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             |||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACGGAAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
```

```
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451

Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511

Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571

Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582

Db       2754 TTGCAACTGCA 2764


RESULT 17
US-10-717-580-14
; Sequence 14, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 2575
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-14

  Query Match             28.6%;  Score 2141.4;  DB 1;  Length 2575;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches 2275;  Conservative  134;  Mismatches  159;  Indels   20;  Gaps   11;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTG-AGYCARTYMTCATAC 2969

Db          1 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAAGCCAATTCTCATAC 60

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029

Db         61 CTGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACC 120

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089

Db        121 TTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTT 180

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149

Db        181 TCCAAACCAAAGACCCAGCTCTGCTCACAGCAGGTTAAATACTTGGGGCTAAAATTATCC 240

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209

Db        241 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCAAGCCTATACTGGCTTATCCTCATCCC 300

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269

Db        301 CAAATCCTAAAGCAACTAAGAGAGTTCCTTAGCATAACAGGTTTCTGCTGAATATGGATT 360

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329

Db        361 -CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGC 419

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db        420 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAG 477

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db        478 AGGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTAT 537

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db        538 GTCACAGAAAAAACAGGAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTT 597

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db        598 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 657

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db        658 TTTATGGGTAGTGGCGGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATACAGGG 717

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689

Db        718 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 777

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749

Db        778 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGT 837

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809

Db        838 GCAGCAACTGCGCAGTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATG---- 893

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869

Db        894 -AAGATAGAACATAACTGCCAACAAGTAATTTCTCAAACCTAGGCCGCTCGAGGGAACCT 952

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929

Db        953 TTTAGAGGTTCCCTTAACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 1011

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989

Db       1012 TAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGCTCAGTGATAATGGAATACTTGAAA 1071

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049

Db       1072 ATAATCCCTTCATTCCAGGAACCAGCGTTCAGCTGGCAGAATTAATAGCCCTCACTCGGG 1131

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109

Db       1132 CATTAGAATTAGGAGAAGGAAAAAGGGTAAATACACATACAGATTCTAAGTATGTTTACT 1191

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169

Db       1192 TAGTCCTCCGTGCCCACGCAGCAATATGGAGAGAAAGGGAATGCTTAACTTCTGAGGGAA 1251

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229

Db       1252 CACCTATCAAACATCAGGAAGTTATTAGGAGATTATTATTGGCTATACAGAAACCTAAAG 1311

Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289

Db       1312 AGGTGGCAGTCTTACACTGCTGGGGTGGTCAGAAAGAAAAGGAAAGGGAAATAAAAGGGA 1371
```

```
Qy       4290 AYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGC 4349
             |:|||||||| :|||||||||:|||||||| |||||||||||||||||||||||||||||
Db       1372 ACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGC 1431

Qy       4350 TTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAG 4409
             ||||  ||   ||| |||||| |||||||||| |||||||||||||||||| |||| ||
Db       1432 TTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAATACTTAG 1491

Qy       4410 CAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAG 4466
              |  ||||| |||||||||||||||||||    ||||| |||||||| ||| || |||
Db       1492 AAAAAGAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCTCCCCTCAGGATGGCTAG 1551

Qy       4467 CCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCC 4526
             |||| |||||||| |||||||||||||||||| ||| ||||||||||||||||||||||
Db       1552 CCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCC 1611

Qy       4527 TTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTA 4586
             |||| |||||||||||||||| ||| |||||||||||||:||||||||||||||||||||
Db       1612 TTCACCAAACCTTTCACTTAGACATTGATAGCACCCATCAGATGGCCAAATCATTATTTA 1671

Qy       4587 CTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARA 4646
             ||||||||||||||||||||||||||||||: ||:|||||||||||||:|||||| :|
Db       1672 CTGGACCAGGCCTTTTCAAAACTATCAAGCAGCTAGTCAGGGCCTGTGAAGTGTGCCGAA 1731

Qy       4647 RAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACC 4706
              :|||||||||| |||||:||| |||||||||||||||||:||||||:|||||||||||||
Db       1732 GAAATAATCCCATGCCTTATCACCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACC 1791

Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
              | |:|:||:::|||||||| |||||||||||||| ||||||| ||||||||||||||||||
Db       1792 CAGGAGAAGRVTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 1851

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
              ||||||| ||||||:|||||||||| |||||||||:|| |||||||||||| ||||||||
Db       1852 TACTAGTTTGGGTAGATACTTTCACTGGTTGGGCAGAGACCTTCCCCTGTAAGACAGAAA 1911

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
              || |||||||||||||||||||| |||||||||||||||||||||||||||| ||||||||
Db       1912 AGTCCCAAGAGGTAATAAAGGCATTAGTTCATGAAATAATTCCCAGATTCAGACTTCCCT 1971

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
              |||||||||||||||||| |||||||||| |||| |||||||||| ||||||||||
Db       1972 GAGGCTTACAGAGTGACAATGGCCCTGCTTTCAAGGCTACAGTAACCCA-GGAGTATCCC 2030

Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
              ||| |||||||||| |||||||||||||||||||| |||   ||||||||||||||||
Db       2031 AGGTGTTAGGTATACAATATCACTTACACTGCGCCTGGAGG---CAGTCCTCAGGGAAGG 2087

Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCT-AAAAAAGCAAACCCAGGAAACCCAC 5064
              ||||||| |||||||||:|||||| |||||||| |||||||||| |||| ||||
Db       2088 CCGAGAAACTGAATGAAACACTCAAACGACATCTAAAAAAAGCTAACCCAGGAAAACCAC 2147

Qy       5065 CTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAA 5124
              ||||||||||||:|||||||||||||||||||| |||||||   |||| ||||||||||
Db       2148 CTCACATGGCCTGCTCTGTTGCCTATAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAA 2207

Qy       5125 GCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 5184
              |||||||||||||||||||||||||| |||||| ||||||| ||||||||||||||||||
Db       2208 GCAGGACTTAGCCCATACGAAATGCTATATGGATAGCCCTTCCTAACCAATGACCTTGTG 2267

Qy       5185 CTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAA 5243
              |||||| | ||| |||||||||||||||||||||||||||||||||| ||||||||||||
Db       2268 CTTGACTGAGAGAGAGCCAACTTAGTTGCAGACATCACCTCCTTATCCAAATATCAACAA 2327

Qy       5244 GTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCW 5303
              |||||||||||||||||||| ||| |||| |||||||||| || ||||||||||||||:
Db       2328 GTTCTTAAAACATTACAAGGAGCCTGTCCCCGAGAAGAGGGGAAGGAACTATTCCACCCT 2387

Qy       5304 WGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGG 5363
              |||||||||||||||||||||||||||:||||||||| |  | |||||||| |||||||
Db       2388 GGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCTCATTGCCTAGATATATCCTGG 2447

Qy       5364 GAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAG 5423
              |||||||||||||||||||||||||:|||||||| || || ||||||||||||||||||
Db       2448 GAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGCAGTAAAAGTGGCTGGAGTGGAG 2507

Qy       5424 TCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCA 5483
              |||||||||||||||||||| |||||||| |||||||| | ||||||||||||||||||||
Db       2508 TCTTGGATACATCACACTCGAGTCAAACCCTGGATATTACCAAAGGAACCTGAAAATCCA 2567

Qy       5484 GGAGACAA 5491
              ||||||||
Db       2568 GGAGACAA 2575
```

RESULT 18
US-10-717-580-13
; Sequence 13, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISF
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 2585
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (726)..(726)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (833)..(833)
; OTHER INFORMATION: n = any nucleotide
US-10-717-580-13

Query Match 27.2%; Score 2036.1; DB 1; Length 2585;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2233; Conservative 129; Mismatches 213; Indels 19; Gaps 14;

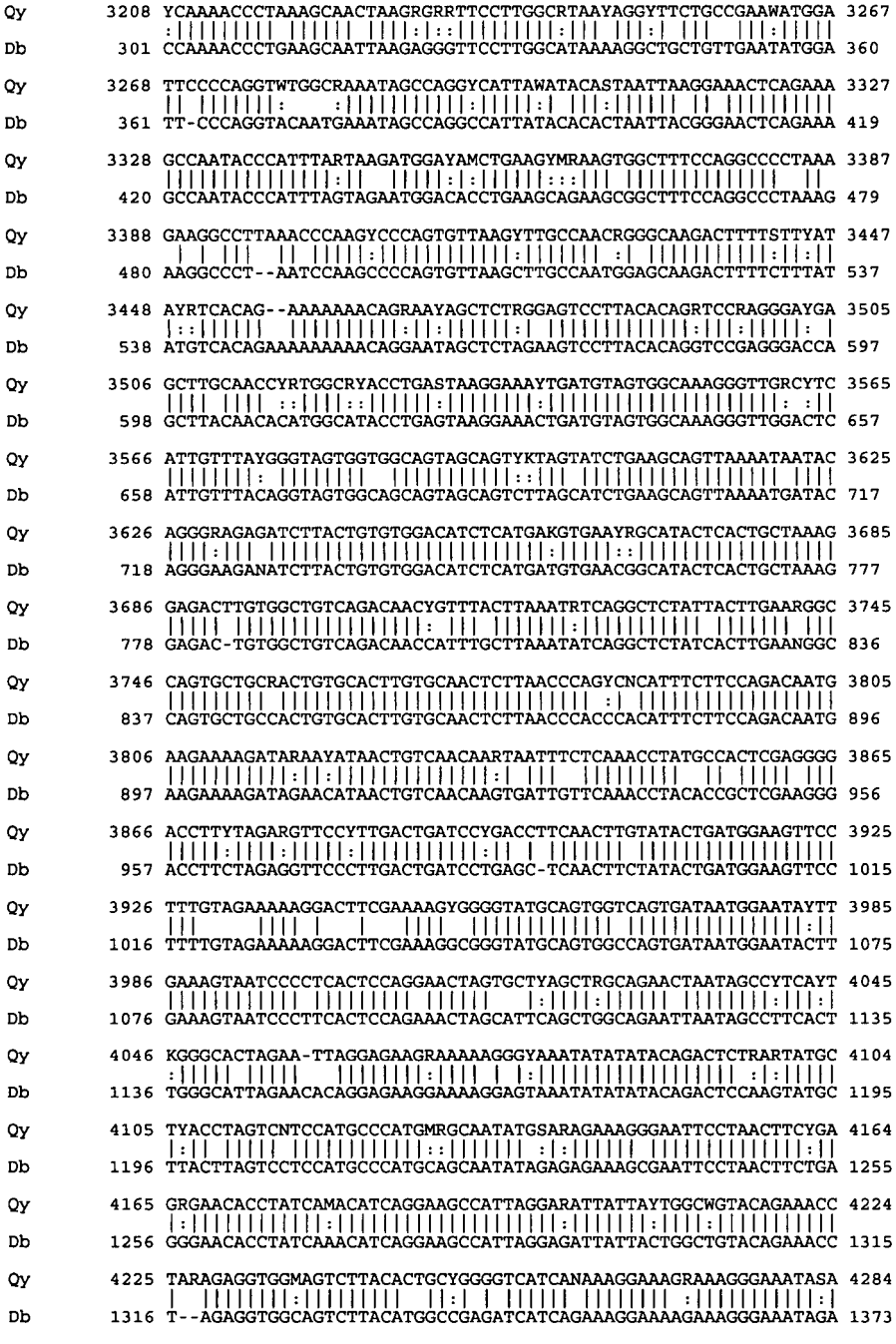
```
Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
              ||||||||||:||||||||||||||||||| |||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

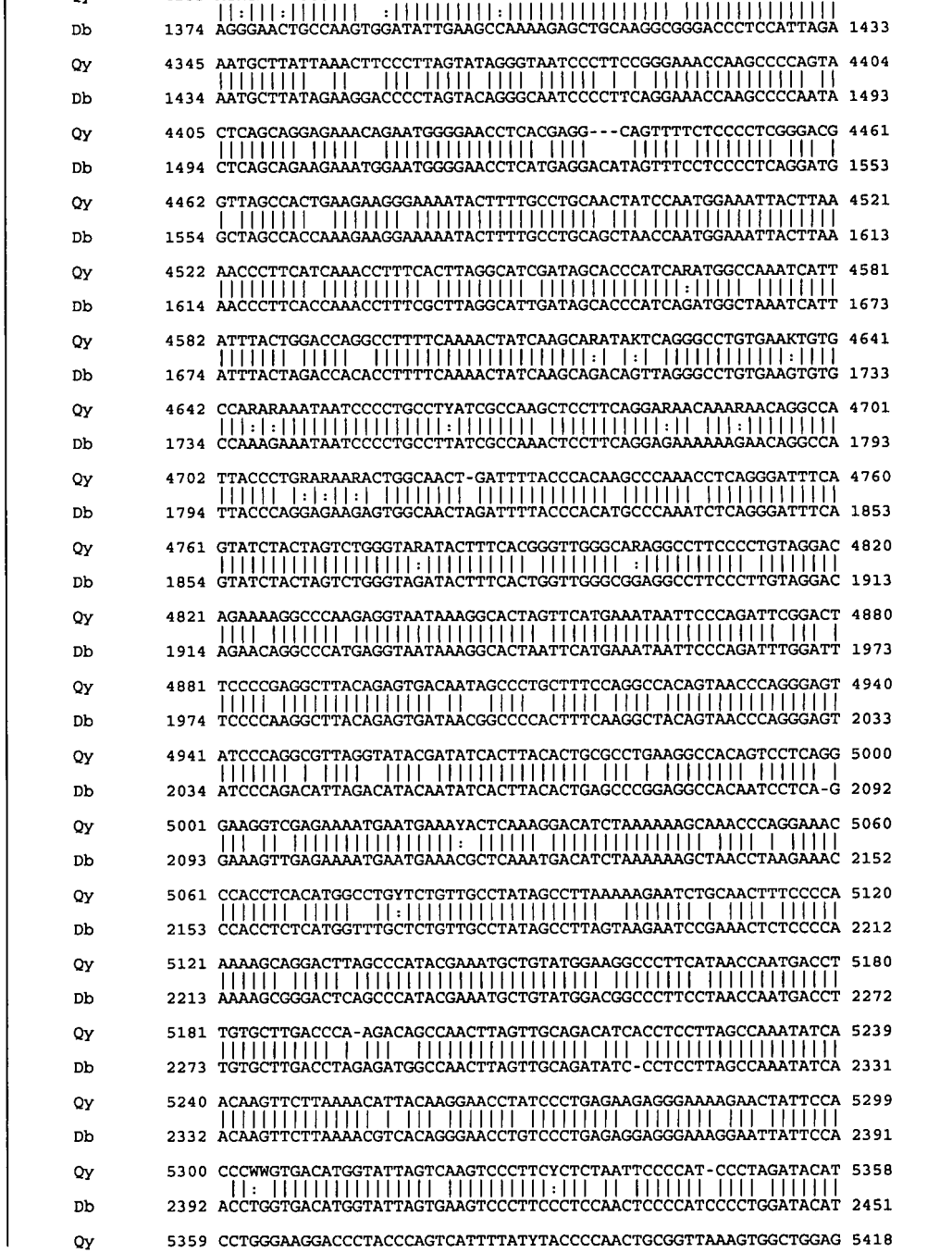
Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
              || |||||||||||||||||| :|||:|||||||||| |||||||:||: ||::||||||||
Db         61 ACTTGGACACTCTTGTCCTTTGGTATGTGGATGATCTACTTTTAGCCACCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
              |||||||||||||||||| ||||| :|||||:|| |||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCAACCAAGTGCTCTTAAACTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
              |||||||||:|:|:|||||:|||||||| |||||||||||||||||||:||||:|||||
Db        181 TTTCCAAACCAGAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
              ||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300
```





```
                ||||||||||||||||| |||||||||||:|| |||||| ||||||||| |||||||||
Db       2452 CCTGGGAAGGACCCTACTCAGTCATTTTATCTATCCCAACCGCGGTTAAAATGGCTGGAG 2511

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCA-AAGGAACCTGAA 5477
                | || |||||||||||||||| | |||||||| ||| |||||||||| ||||||||||
Db       2512 TAGAATCTTGGATACATCACATTCGAGTCAAACCCTAGATACTGCCACAAGGAACCTGAA 2571

Qy       5478 AATCCAGGAGACAA 5491
                ||||||||||||||
Db       2572 AATCCAGGAGACAA 2585


RESULT 19
US-11-028-539-22
; Sequence 22, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/11/028,539
;  CURRENT FILING DATE:  2005-01-05
;  PRIOR APPLICATION NUMBER: US/09/719,554
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 2055
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2055)
;   OTHER INFORMATION:
US-11-028-539-22

  Query Match             27.1%;  Score 2032.5;  DB 1;  Length 2055;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 2025;  Conservative   30;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
                ||||||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db         61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
                |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
                |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
                |||||||||||||||||||||||| ||||||||:||||||||||||||||||||||:
Db        361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
                ||:||||:|||:||||||:||||||||||||||||||||||||||||||||||||||||
Db        420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
                ||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db        480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
                |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
                ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
                |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
                ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
                ||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
```

```
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809

Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869

Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929

Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989

Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049

Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109

Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169

Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229

Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245

Db       2040 TAGGTTTTCCTGTTGA 2055


RESULT 20
US-08-979-847-87
; Sequence 87, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  87:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2304 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-979-847-87

  Query Match             24.7%;  Score 1851.9;  DB 1;  Length 2304;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches 2032;  Conservative  119;  Mismatches  150;  Indels   29;  Gaps   15;

Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084

Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141

Db         59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201

Db         119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261

Db         179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321

Db         234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381

Db         294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441

Db         354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501

Db         413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561

Db         472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621

Db         532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681

Db         592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741

Db         644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801

Db         703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
```

```
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921

Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981

Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041

Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101

Db       1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161

Db       1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221

Db       1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281

Db       1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341

Db       1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401

Db       1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461

Db       1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521

Db       1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581

Db       1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641

Db       1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701

Db       1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761

Db       1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821

Db       1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881

Db       1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941

Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001

Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061

Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121

Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181

Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241

Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301

Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351

Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 21
US-10-114-104-87
; Sequence 87, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
```

;  TELEPHONE: 703-836-6400
;  TELEFAX: 703-836-2787
;  INFORMATION FOR SEQ ID NO: 87:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 2304 base pairs
;  TYPE: nucleotide
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  SEQUENCE DESCRIPTION: SEQ ID NO: 87:
US-10-114-104-87

Query Match 24.7%; Score 1851.9; DB 1; Length 2304;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2032; Conservative 119; Mismatches 150; Indels 29; Gaps 15;

```
Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             ||||||||||||||||||||||| |||||||| |||| ||||||||||||||||   |||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| ||||||| |||||||||||| |||||||||  |    |||||||||||||||| |
Db         59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             | | ||||| |||||||||  | |||||||  ||||| |||||||||||||||||||||| |
Db        119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                | |||   | ||  ||| | |||||||||| |||| ||:|| ||||||||||||||| || |
Db        179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
             ||||||| ||||| |||||||||||||||||||||||||| ||||||||||||| ||||
Db        234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||| ||||
Db        294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
             || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
             |||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db        413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
               | |||||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db        472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
             |||||||||||||||| |||||||||||||||| |||||||||||||||   |||||||| |
Db        532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
             | | |      |||   | | |  |||||||| |||||||||| ||||||||||||
Db        592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
             |||||||||||||| ||| ||||| ||| | | | |||||||||||||| ||||||||||||
Db        644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
             ||| ||::||||| |||||||||||||||||||||||||||:||||||||||||:|||||
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||| ||
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
               |  ||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:|||

Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
             ||||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
             |||||||:||| ||||||||||||||||||:| :|||::|||||||||||||||||||||
Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
             |||||||||| :|||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:
Db       1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
             |||| :||||||||||||| || ||| |||||||||:||||:||||||||||||||||:|
Db       1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
             |||||||||||||||||:::|||||||||||||||||||||||||||||||:|||||||||
Db       1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
             ||||||||:|::|||||||||:|||:|||:||||||||||:|  |||| |||||||:
Db       1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
             |  |||||||||  :|||||:|||||:|||||| ||||||||||||||||||||| ||||
Db       1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
             |:|||||||||:|:||    :::|||||||||||||||||||  ||  | | |||  ||||
Db       1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
             ||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: ||||| ||||
Db       1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
             |||||:||:||||||:|||||||||||| |||:||  |||||:|||||||||||::|||
Db       1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
             ::||||||:|||||||:|||||||||||||||||||||||:|:|||||||||:|||||| |
Db       1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
             || ||||||||||||::|||||||||||||||||||||||||||||||:||||||||||||
Db       1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
             |||||||||||||||||:|||||::|||||||||||||||||||||||||||||||||||
Db       1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
             |||||||: |||||||||  |:|||||||||||||||||: |||||||||| :|||| |
Db       1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
             |||||||||||||||||  || |:| ||||||||||||||||||||||||||||:||:|
Db       1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
             ||||||||||||:||||| ||||||||||||  ||||||||||||||:||||:|||||
Db       1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
             :|||||||||||:|||| ||||||||||||||||||||||||||||| | ||||||||||
Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
             |||||||||:|||||||||||||| ||||||||||||||||:|||||||||||| |||||
Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954
```

```
Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
             ||||||||||||||||||:| ||:||||||||||||||||:|||:|:|||||||||||||||
Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
             ||||||:||||||||:|||||||||| ||||||||:|:|||||||:|||||||| ||||||
Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
             ||||||::||||||||:|:||| ||||||||||||||||:|||:||||||||||||: |
Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
             ||||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:|||||
Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: |||||||||  ||| | ||:||||||||||:|||: ||: ||||||| :
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
             ||||||||||:|||| ||| |||||||||||| ||||||||||||||||
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 22
US-10-632-793-23
; Sequence 23, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
;   LENGTH: 2006
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (305)..(305)
;   OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-23

  Query Match           24.7%;  Score 1848.6;  DB 1;  Length 2006;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1949;  Conservative   56;  Mismatches    1;  Indels  206;  Gaps    3;

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
             ||:||||:|||||||||||||||:|||:|:|||||||||||||||||:||||||||
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
             :|||||||||||||||:|:|||||:|||||||| ||||||||||::|||||||
Db        121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
             |:|:|||||||||||||||||||||||:|||:||||||||||:||||||||||||||||
Db        181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
             |||:|||||||:||||:|||||||||||||:||||||||:|||||||||||||:||||||
Db        241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
             ||||||||||||||:|
Db        301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376

Db        317 ---------------------------------------------------------- 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436

Db        317 ---------------------------------------------------------- 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
                                                     ||||||||||||||||||||
Db        317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
             |:|||:|||||||||||||:||||||||:|:|||||||||||||||:||||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:||||||:|||||||||||||||||:|:||:|||||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
```

```
                ||||||
```

```
Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
             |:|||:|||||||||||||||:||||||||:|:||||||||||||||||:|||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:||||||:||||||||||||||||:|:||:||||||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
             ||||||||||||||||||||||||||::||||||||||||||||||||||||:||||
Db       1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||||||||||||:||:||||:|||:||||||:|||||||||||
Db       1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             ||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||
Db       1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

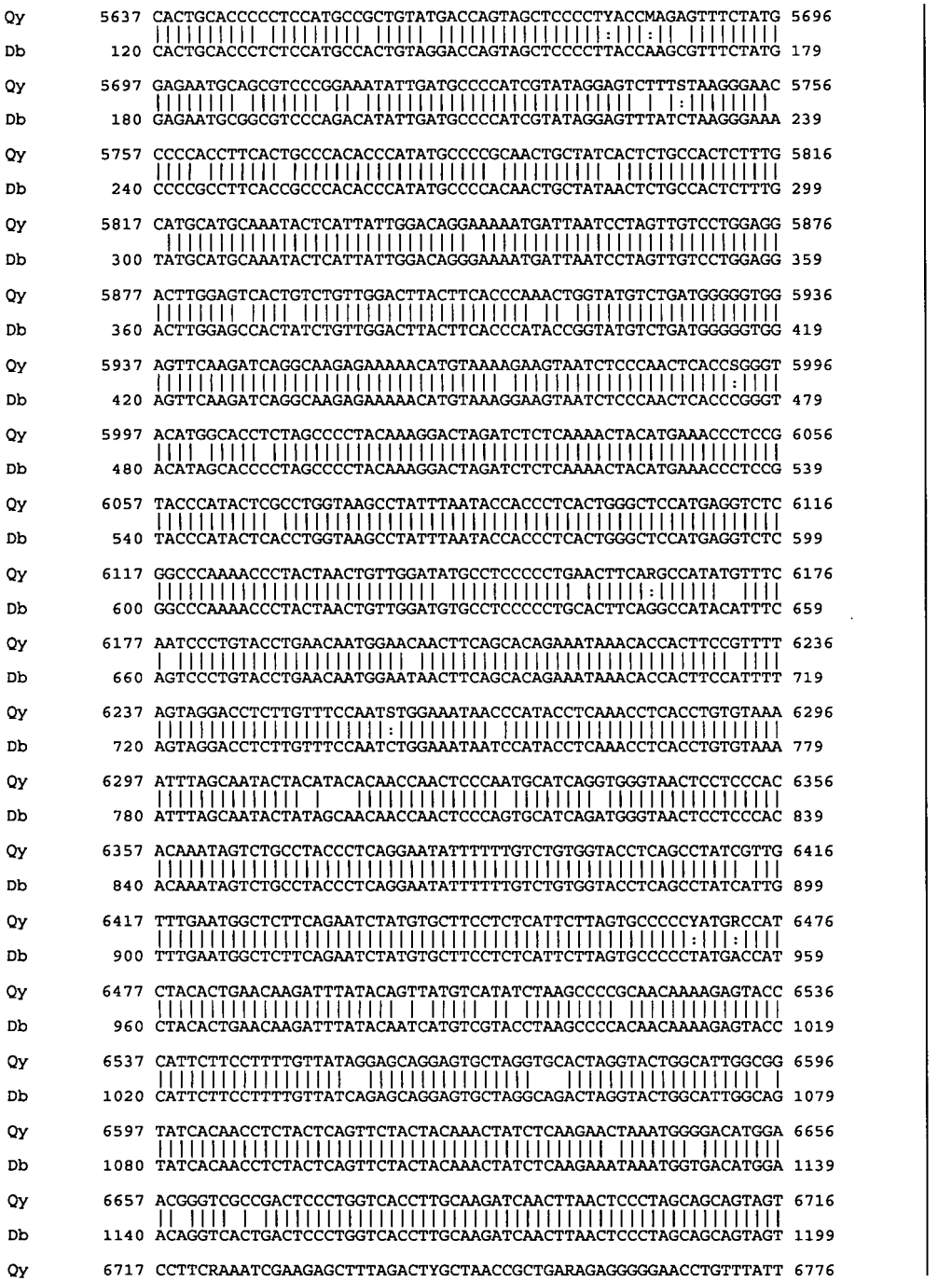
Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006
```

RESULT 24
US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
US-10-416-642-4

Query Match 24.2%; Score 1812.4; DB 1; Length 2074;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1904; Conservative 33; Mismatches 125; Indels 14; Gaps 3;

```
Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             |||||| || ||||||||||||| ||||||||  ||||||||||||||||||| |||||
Db          1 GAGGATCTGTGCCTGCTCTTCAAGCAACAACCGTGAGGAAAGTAACTAAAATCGTAAAT- 59

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||| ||||||||||:||:||||:| |:|||  :||||| |||
Db         60 CCCCATGGCCCTCCCTTATTGTATTTTTCTCTTTACTGTTCTCTCACCACCTTTCAGTCT 119
```



RESULT 25
US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US20040043381A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: PERRON, Herve
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES A
; TITLE OF INVENTION: METHODS FOR DETECTING MSRV-1 RETROVIRUS
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565
; CURRENT FILING DATE: 2003-08-11

```
;  PRIOR APPLICATION NUMBER: US/09/890,340
;  PRIOR FILING DATE: 2001-11-27
;  PRIOR APPLICATION NUMBER: PCT/IB00/00159
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: EP 99420041.8
;  PRIOR FILING DATE: 1999-02-15
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;    LENGTH: 2030
;    TYPE: DNA
;    ORGANISM: MSRV-1 retrovirus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)..(1626)
US-10-637-565-18

  Query Match             23.8%;  Score 1780.4;  DB 1;  Length 2030;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches 1848;  Conservative   33;  Mismatches  119;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db             1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||| ||||||  |||||||||||||||:|||:|||||||||||| |||
Db             61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             | || || ||| |||||||||||||||||||||| ||||||||| | |:||||||||| | ||
Db             121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db             181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db             241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             |||| |||||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db             301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             |||| |||||||||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db             361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             |||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db             421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             |||||||||||||| |||||||||||||||||||||||| |||||||||||||||| |||
Db             481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             |||||||||||||||||||||||| ||||||||||| |||||:|||||| ||||||||
Db             541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db             661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||| | |||||||| |||||||||||||||||||||| |||||||||| |
Db             721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db             781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||
Db             841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             |||||||||||||||||| | ||||| || ||||||||| ||||||||||||||||||||
Db             901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             |||||||||||||||  |||||||||||||||||  |||||||||||||||||| |||||
Db             961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||| |
Db             1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:|||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||||
Db             1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||||| |||| |||||||||||||:|||||
Db             1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:|||||||:||||||| |||||| | || ||||||||||
Db             1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| ||||||||||||||||||||||||||||||| |||||||||
Db             1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
               |||||||||||||||||||||||||:|||||||||||||| |||||||||||||||||
Db             1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAAC------------TACAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
             ||||||||:|| |            |||||||||| ||| ||||| |||||| |||||||
Db             1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCAGATGGAGTCCATGACTAAG 1500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
             ||| |||| |||||||||||||||||:||||||| | || |||||||||||||| ||
Db             1501 ATCCACCGTGGACCCCTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAA 1560

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
             |||||||||| |||||||||| |||||||||| |||||| |||||||||||| |||
Db             1561 GGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGA 1620

Qy       7189 AGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAT 7248
             |||||||||||||||: || |||||||||||||||||||| ||||||||||||||||||
Db             1621 AGCAGTTAGAGCGGTCATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAG 1680

Qy       7249 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTA 7308
             |||||||||||||||||||||||||||||||||||||||| |:|||||||:|||||||||
Db             1681 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTA 1740

Qy       7309 GSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGA 7368
             |:|||||||||||  |||||||| |||||| |||||||||||||:|||||||| ||
Db             1741 GCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGA 1800

Qy       7369 CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCA 7428
             |||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db             1801 CCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCA 1860

Qy       7429 ATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYT 7488
```

```
            |||||:||||||:||||||||||| |||||||||| |||||||||||||||||| |  |
Db       1861 ATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCAT 1920

Qy       7489 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTC 7548
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1921 TCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTC 1980

Qy       7549 ATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
                | |||||||||||||||||||||||||:|||
Db       1981 ACTCTATTTCACTCTATTAAATCTTGCAACTG 2012


RESULT 26
US-10-133-036-3
; Sequence 3, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
;  APPLICANT: Conrad, Bernard
;  APPLICANT:  Mach, Bernard
;  TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
;  FILE REFERENCE: 23135-507
;  CURRENT APPLICATION NUMBER: US/10/133,036
;  CURRENT FILING DATE:  2002-04-26
;  PRIOR APPLICATION NUMBER: PCT/EP00/10659
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1617
;   TYPE: DNA
;   ORGANISM: Human endogenous retrovirus
US-10-133-036-3

  Query Match             21.5%;  Score 1607;  DB 1;  Length 1617;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1592;  Conservative   25;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||:||:||||:|||:||||||:||||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:|||||||||||||||||||||:||||||||||||:|||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db       1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 1617
```

```
RESULT 27
US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4444
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC002346.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  20
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;
```

```
Qy      5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db         1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy      5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db        61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy      5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||||             ||||| ||||||| || ||  ||||||
Db       121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy      5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db       168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy      5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||
Db       228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy      6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| ||||||||||||||||||||||||||||||| |||||||
Db       288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy      6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db       348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy      6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy      6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||| ||||   |||||||:|||||||||||||||||||||||||||
Db       468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy      6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||||
Db       525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy      6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy      6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy      6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||| | |||||||| ||||||||||||
Db       705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy      6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db       765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy      6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy      6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||| ||||||| ||  |||||||||||||||||||||||||||||||||
Db       885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy      6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:||||||||| ||:|| ||||
Db       945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy      6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| ||||||||| ||||||||||||||||||||||||| |||
Db      1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064
```

```
Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||||| || ||||||| ||||||:|||| |:|||||  | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| |||||||||||||||||||||||:||||||||||||||  |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051
             ||||||  |||||||||||||||  ||:||||||          |||||||||||  |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             |||||||||  ||  |||  |||||||  | ||||||||||||||||||:|||||||  | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||| ||| || ||||||||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| |||||||||||||||| | |||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||          ||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||          ||||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 28
US-11-028-539-5/c
; Sequence 5, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-028-539-5

Query Match 19.9%; Score 1487.1; DB 1; Length 1799;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 1613; Conservative 27; Mismatches 116; Indels 45; Gaps 5;

```
Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||| ||||| ||||||||:|||||||||||||||||||
Db       1799 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 1741

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||||||||||
Db       1740 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 1681

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
                 ||||| |||||||||:|||:|||||||||||| ||||||| || |||| |||||||||
Db       1680 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 1621

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
             |||||||||| |||||||| |  :||| |||| | ||||||||||||||||||||||||||
Db       1620 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 1561

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
             |||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 1501

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
             |||| |||||||||||||||||||||||||||||||| |||||||||| |||||||||| ||||||
Db       1500 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 1441

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 1381

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAG 6022
             ||||||| ||||||| |||||||||| |||:|||||||| ||||| ||||||||||||||||
Db       1380 ATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCTACAAAG 1321

Qy       6023 GACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTAT 6082
             |||||||||||| |||||||||||||||||||||||||||||||| |||||||||||||||
Db       1320 GACTAGATCTCTTAAAACTACATGAAACCCTCCATACCCATACTTGCCTGGTAAGCCTAT 1261

Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGA 6142
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1260 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAAACCCTACTAACTGTTGGT 1201

Qy       6143 TATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACA 6202
             | |||||||||||| | ||||:|||||   |||||||||||||||||||||||||||||||
Db       1200 TGTGCCTCCCCCTGTATTTCAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACA 1141

Qy       6203 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1140 ACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGG 1081

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
             |||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db       1080 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCA 1021

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1020 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAA 961

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        960 TATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 901

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
             |||||||||||||||||||||||||||||| |||:|||| |||||||||||||||||| ||
Db        900 GCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATT 842
```

```
Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
             ||||  || |||||||| |||||||||||||| |||||||||||||||||| |||||||
Db        841 ATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAG 782

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
             |||||||||||| ||||||| |||||||||||||| |||||||||||||||||||||||
Db        781 GAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACT 722

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
             |||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||||||
Db        721 ACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCT 662

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
             |||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||||||
Db        661 TGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTAGACT 602

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGG---------- 6782
             :||||||| ||||:|||||||||| ||||||||||||||
Db        601 TGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATG 542

Qy       6783 ----------GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAG 6832
                      ||||||||| |||||||||||||||||||||| |||| ||||| |||||||
Db        541 TTATTTTAGCGGAAGAATGTTGTTATTATGTTAATCAATCCTGAATTGTCACAGAGAAAG 482

Qy       6833 TTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC-GAAACACTGGACCC 6891
             ||:|||||||||:|||| ||||||||||||:| ||||:|||||||  ||||||  |||||
Db        481 TTGAAGAAATTCGAGATTGAATACAACGTAGAACAGAGGAGCTTCAAAAACACCAGACCC 422

Qy       6892 TGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCT 6951
             |||||||||||||||:|||||||||||||||||||||||||||||||| ||||||||||
Db        421 TGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGATCTCTAGCAGCT 362

Qy       6952 ATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCT 7011
              ||||||||| ||||||||||||||||||||||||:|||||||||||||| |||||||||
Db        361 CTAATATTGATACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCT 302

Qy       7012 TCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCC 7059
             ||||||||| ||| |||:|| |||            |||||||| ||| ||||| |||||
Db        301 TCCAGAATCAAAGTTGTAAAGCTACAAATCGTTCTTCAAATGGAACCCCAGATGAAGTCC 242

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             | ||||||||||||||  |||||||||||||||||| :|||||| | ||  ||  ||||
Db        241 ATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTACTAGCCCATGCTCCAATTGTAAT 182

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             || ||| || |||||||||| |||||||||||| |||||||||| |||||| |||||||||
Db        181 GATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAAT 122

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             || ||||||||||||||||  |||:||| |||||||||||||||||||| |||||||||
Db        121 TCCGCAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCCAACAGCACTTGGGTTTTCC 62

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||  |||||||||||||||||| |||||||||||||||||| ||:|||||||||
Db          61 TGTTGAGTGGGGGGACTGAGAGACAGGATTAGCTGGATTTCCTAGGCCGACTAAGAATCC 2

Qy       7300 Y 7300
              :
Db          1 C 1


RESULT 29
US-10-632-793-24
; Sequence 24, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (84)..(84)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (193)..(193)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (241)..(241)
;   OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-24

  Query Match           19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db         491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db         671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             |||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db         730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db         790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCCACCTTCACTGC 5771
             |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db         850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
```

```
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
             ||||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948


RESULT 30
US-10-717-580-5
; Sequence 5, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;  LENGTH: 1948
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (84)..(84)
;  OTHER INFORMATION: n = any nucleotide
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (193)..(193)
;  OTHER INFORMATION: n = any nucleotide
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (241)..(241)
;  OTHER INFORMATION: n = any nucleotide
US-10-717-580-5

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
```

```
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
                |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
                |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
                |||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
                |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
                ||||||||||:||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
                ||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948


RESULT 31
US-10-450-763-9505
; Sequence 9505, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 9505
;   LENGTH: 1715
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (668)..(868)
;   OTHER INFORMATION: 83% homologous to Homo sapiens Protein encoded by the
;   OTHER INFORMATION: endogenetic fragment of HERV-W,accession number B07702,Smith-
;   OTHER INFORMATION: Waterman Score=300.
US-10-450-763-9505

  Query Match             18.5%;  Score 1383.7;  DB 1;  Length 1715;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches 1551;  Conservative   29;  Mismatches   97;  Indels   39;  Gaps   11;

Qy         90 CTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTATTTCACTCTATTAAATCTTGCAACTGCATTCTTCTGGTCCGTGTTTGTTACAGCTCG 60

Qy        150 AGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTT-----------------GCCACCA 192
              ||||||||||  ||  | | |||||||| ||||                  ||| |
Db         61 AGCTGAGCTTTCACTTGCTGCCCACCACTGGTGTTTGCCGCTGTCACAGACCCGCCGCTG 120

Qy        193 CCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCT 252
              |||| ||| || |||||||| |||||| ||||||||| |||||||||||| ||||||||||
Db        121 TCGCAGACCCGCTGCTGACTGCCATCCGTCTGGATCCGGCAGGGTGTCCACTGTGCTCCT 180

Qy        253 GATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGC 312
              ||||||  ||:|| |||||||||||||||| ||  |||||||||||||||||||| |||||
Db        181 GATCCAATGAGGCACCCATTGCCGCTCCCGATCCGGCTAAAGGCTTGCCATTGTTCCTGC 240

Qy        313 ACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGT 372
              || |||| ||||| |||||||| |||||  |||||||||||| |  |||||||||| |||
Db        241 ACAGCTAGGTGCCCGGGTTTGTCCTAATCAAGCTGAACACTACTTGCTGGGTTCCACGGT 300

Qy        373 TCTCTTCTGTGACCCACGGCTTCTAAT---------AKAACTATAACACTTACCACATGG 423
              ||||||| ||||||||||||||||||         |:| |||||||||| ||| || ||
Db        301 TCTCTTCCGTGACCCACGGCTTCTAATAGAGCTAACAGAGCTATAACACTCACCGCACGG 360

Qy        424 CCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGAR 483
              |||||||||||||||||| ||||||||||:|||  | | ||||||||||||||||||:
Db        361 CCCAAGATTCCATTCCTTCGAATCCGTGAGGCCAAGAATCCCCAGGTCAGAGAATACGAA 420

Qy        484 GCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTG 543
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        421 GCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTG 480

Qy        544 GGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCMAA 603
              |||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||:||
Db        481 GGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCCAAA 540

Qy        604 GTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAATTGG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        541 GTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAATTGG 600

Qy        664 GAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCC 723
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCCGCC 660

Qy        724 TGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||:
Db        661 TGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTGTAG 720

Qy        784 AAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 780
```

```
Qy       844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
Db       781 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 840

Qy       904 TCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCC 963
Db       841 TCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGGTCC 900

Qy       964 AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTAT 1023
Db       901 AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTAT 960

Qy      1024 GACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTT 1083
Db       961 GACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCTTTT 1019

Qy      1084 YYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGAT 1143
Db      1020 TCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCTGAT 1079

Qy      1144 GGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA 1203
Db      1080 GGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATA 1139

Qy      1204 TATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGC 1263
Db      1140 TATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGC 1199

Qy      1264 CTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAA 1323
Db      1200 CCGAGAGCTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-A 1256

Qy      1324 GGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGG 1383
Db      1257 GGAAAGAGAACGATTCCCTACAGGCCAGCAGGCAGTTCCCAGTGCAGACCCTCATT---G 1313

Qy      1384 ACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTAS 1443
Db      1314 AGACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTACTAACTTGCGTGCTGG 1372

Qy      1444 AAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACA 1503
Db      1373 AAGGACTAAGGAAAACCAGGAAG-AAGGCTGTGAATTTTTCAGTGATGTCCACTGTAACA 1431

Qy      1504 CAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAA 1563
Db      1432 CA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCACTGAGGAA 1490

Qy      1564 GCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTAT 1623
Db      1491 GCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGATTAAGTTTAT 1550

Qy      1624 CACTCAGTCAGCTGCAGACATTAGAAA--AAACTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
Db      1551 CACTCGGTCAGCTGCAGACATTAGAAAACAAACTTCAAAAGTCCACCTTAGGCCTGGAGC 1610

Qy      1682 AAAACTT-AGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAG 1740
Db      1611 AAAACTTAAGACACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAG 1670

Qy      1741 GAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA 1776
Db      1671 GAGCAGTCGGAATGGGACAAACGAGATTAAAAAAAA 1706


RESULT 32
US-10-632-793-22
; Sequence 22, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 1422
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (879)..(879)
;   OTHER INFORMATION: n = a or g or c or t/u
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1200)..(1200)
;   OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-22

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy      2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
Db         1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy      2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
Db        61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy      3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
Db       121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy      3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db       181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy      3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db       240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy      3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db       300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy      3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db       360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy      3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db       420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy      3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db       480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy      3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
Db       540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy      3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
```

```
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||||||||||||||||||||:|||||::||||||||||||||||||||||
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||:||||||||||||:||||||||||||||||||||:|||||
Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||||||||||||||||||||||||||:||||||||||||||||||||
Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
                 ||||||:||:|||||||||||||:|||||||||||||||||||||||||||||||||||
Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:||||||:||||||||||||:||||||||||||||||||||||||||||||||
Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
             |||||||||||||||| |||||:||||||||||||||||||||||||||||||:||||
Db        1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
             |||| ||||||||||||||||||||||||||:||||:|||||||||||||||:|||:|:|
Db        1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
             |||||||||||||||||||:||||||||:|||||||||||||||||||:|:|||||||:|
Db        1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
             ||||||||||||||||||::||||||||:|:|||||||||||||||||||||:||||:|
Db        1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227
             ||||||||||:|||||||||||||||||||||||||:|||||||:||||:|||||||||:
Db        1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287
             ||||||||:|||||||||||||:|||||||||| ||||||||:||||||||||||:|||:
Db        1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337
             |||:||||||||||:||||||||||:||||||||||||||||||||||||
Db        1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 33
US-10-717-580-3
; Sequence 3, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1422
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (879)..(879)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1200)..(1200)
;   OTHER INFORMATION: n = any nucleotide
US-10-717-580-3

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||:||||||||||||||||:||:|||::|||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||:|||||:||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||| :|:|:|||||:|||||||||||||||||||||||||||||:||||:|||||
Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||:::|||||||||||||||||||||||||||||||||
Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|||||
Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             ||||||||||:||||:|||||||||:|||||:|||||||||||||||||||||||||||
Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||||||||:||||||||:|:||||||:::|||||||||||||||||||||||
Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             ||||||||||||||||||:|||||||||||||:||||||||:|||||||||||||:||:||
Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |::|||||||||||||||:||:||||||:|||||||||||||||:|||:||||||:||||
Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
             ||||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
```

```
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807

Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867

Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927

Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987

Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047

Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107

Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167

Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227

Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287

Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337

Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 34
US-08-979-847-105
; Sequence 105, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  105:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1481 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-105

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640

Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700

Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760

Db         121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820

Db         181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880

Db         241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940

Db         301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000

Db         361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060

Db         421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120

Db         481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180

Db         541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240

Db         601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660
```

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||| | |||||||||| |||||||||||||||||||||||| |||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             |||||||||||||||||| | ||||| || |||||||||| ||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             |||||||||||||  ||||||||||||||||   ||||||||||||||||||| |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||||||| ||||||| |||||||| |
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||||| |||| |||||||||||||||:||||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:|||||||:||||||| |||||| | || ||||||||||
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:||||||||||||| |||||||||||||||||||||||||||| ||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||:|||||||||||| ||||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
             ||||||||:|| ||||| ||||      | ||| |  || |
Db       1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481


RESULT 35
US-10-114-104-105
; Sequence 105, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 105:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1481 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-10-114-104-105

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             |||||||||||||| ||||||  |||||||||||||:|||:|||||||||||| |||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             | || || ||| |||||||||||||||||||||| ||||||||| | |:||||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

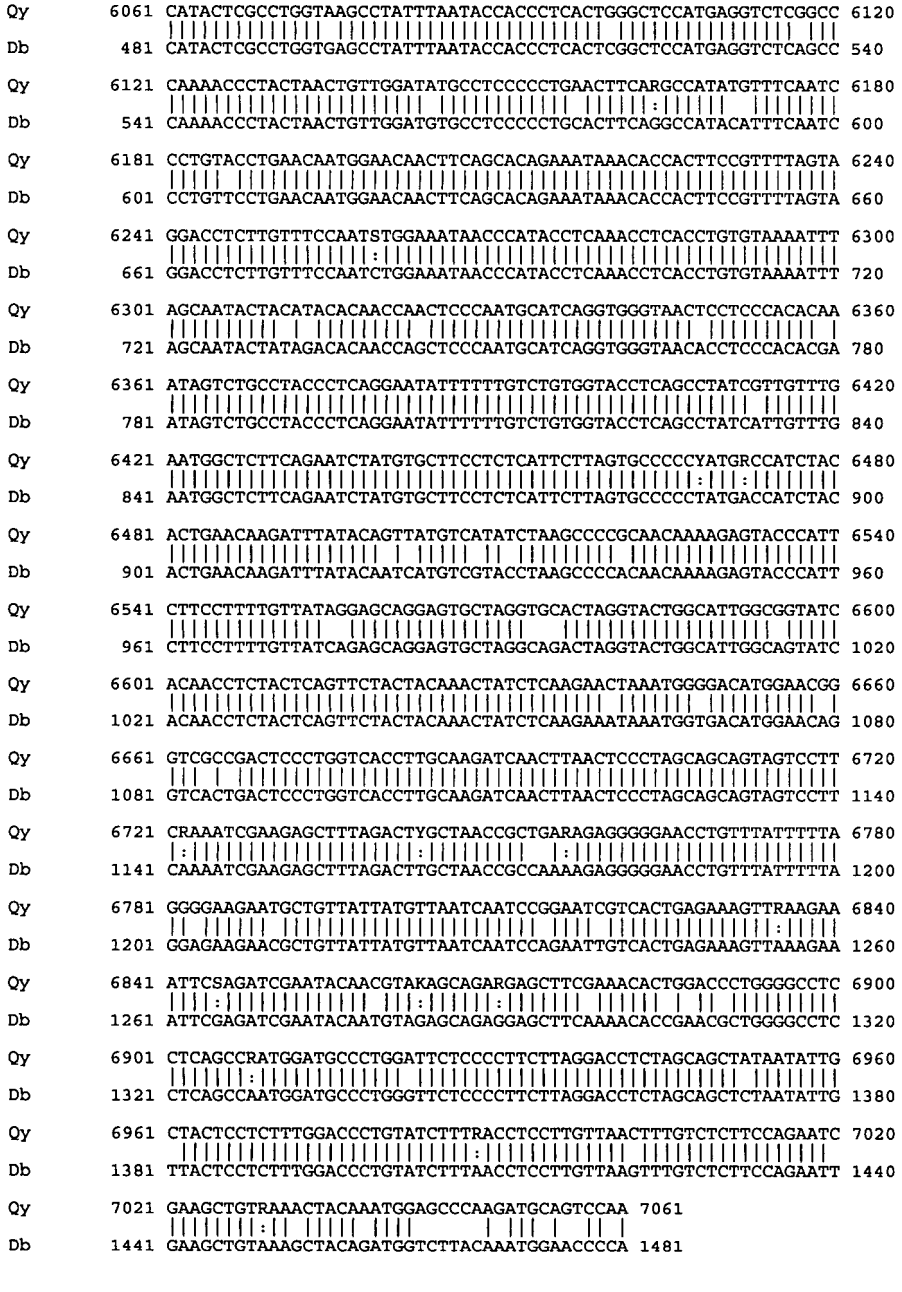
Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             |||| ||||||||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             |||| |||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

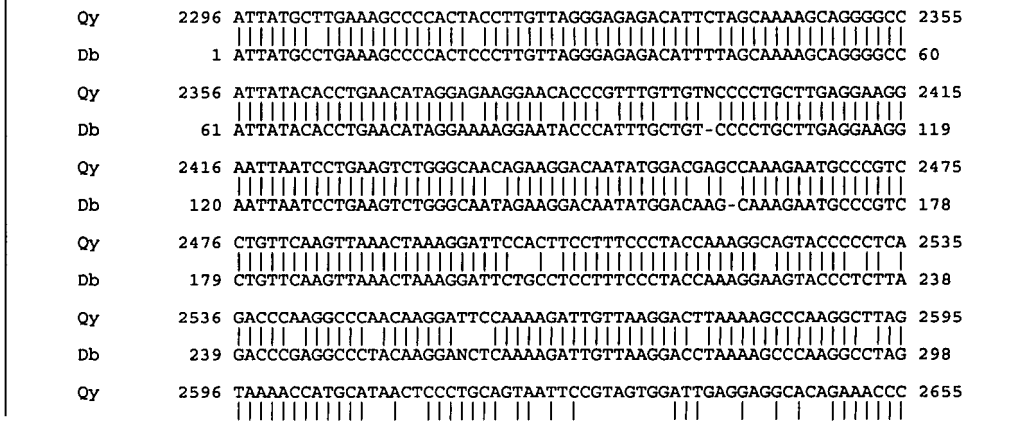
Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480
```



RESULT 36
US-08-979-847-207
; Sequence 207, Application US/08979847

; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-979-847-207

Query Match 17.6%; Score 1320.4; DB 1; Length 1600;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1402; Conservative 111; Mismatches 87; Indels 16; Gaps 9;



```
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715

Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775

Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835

Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895

Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955

Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015

Db        650 GYCARTYYTCATACCTGGACAYTCTTGTYCYTCRGTACRTGGATGATTTACTTTTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075

Db        710 CCCRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGMACTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135

Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCRGCTCTGCTCACAGSAGRTTARATACTTAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195

Db        830 GGCTAAAATTATCCAAAGGCACCAGRRCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255

Db        890 STTATCCTCATCCCAAAACCCTAAAGCAACTAASAGSGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315

Db        950 GCCRAAWAYRGATT-CCCAGGTACASCMMRRTAGCCAGACCATTAWATACACKAATTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375

Db       1009 GAAACTCARAAAGCCARTACCYATTTAGTAAGATGGACAYCTGAAGCAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435

Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495

Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555

Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615

Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675

Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735

Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795

Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855

Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

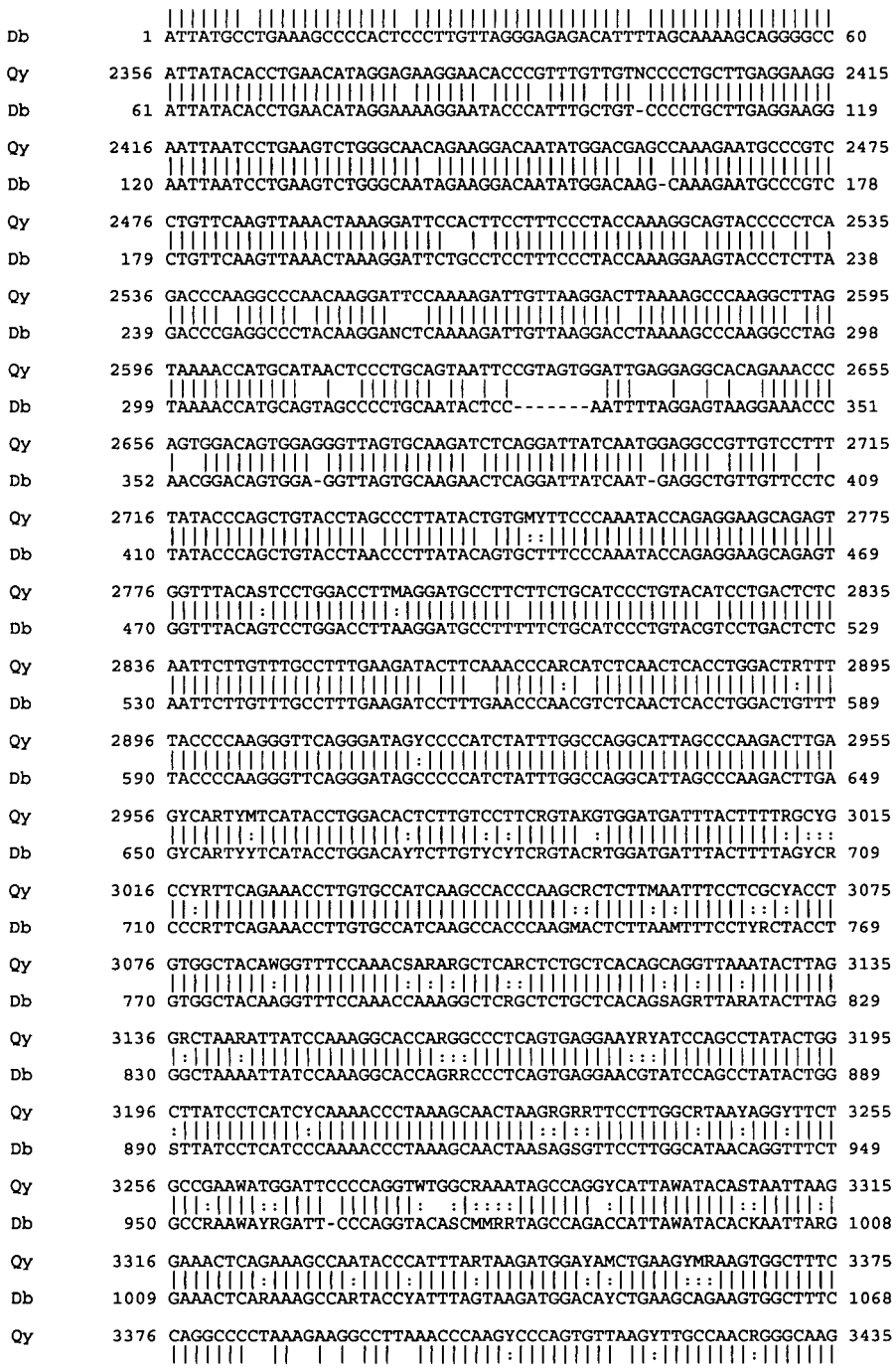
Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911

Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 37
US-10-114-104-207
; Sequence 207, Application US/10114104
; Publication No. US20030198647A1
;  GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/114,104
;            FILING DATE: 03-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/979,847
;            FILING DATE: 26-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;  INFORMATION FOR SEQ ID NO: 207:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1600 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-10-114-104-207

  Query Match            17.6%;  Score 1320.4;  DB 1;  Length 1600;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches 1402;  Conservative  111;  Mismatches   87;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
```




RESULT 38
US-08-979-847-206
; Sequence 206, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  206:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1600 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-979-847-206

  Query Match             17.5%;  Score 1309.8;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||| |||||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||| | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| ||||||| |||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |         |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||| ||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| |||  ||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||:||||||:|:||:|||:::|||||||:|||:||:|:|:::
Db        650 GYCARTYCTCATACCTGGACAYTCTTGTYCYTCAGTAYRKGGATGAYTTAMTTWTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||:|| |||||||||||:|||||:|:||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGSCAMCAAGCCACCCAAGHRCTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:::||||::|||||||||||::||:|||:|||||| ||
Db        770 GTGGCTACAAGGTTTCCAAACMAARGGCTCRSCTCTGCTCACASSAGRTTARATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||:|:|||:|||||||||:||||||:::|||||||:||||||||
Db        830 GGCTAAAATTATCCAAAGKCRCCAGGGCCCTCAGWGAGGAACGTATCCAGCSTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||| ||||:||:|||||:|||||||||||||:::|||||||||:||::||:|||||
Db        890 MTTATCCMCATCCCAWAACCMTAAAGCAACTAAGARGGTTCCTTGGCATAWCAGSYTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||||::|||| |||:|:|:  ::::|||||||||::|||||::||||:||::||:|
Db        950 GCCGAAWAYRGATT-CCCVGRTACASYSMAATAGCCAGRCCATTATRTACAYTADYTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||||:||:||:|||||||||:|:||||: :::|||||||||||
Db       1009 GAAACTCAGAAAGCCAATACCYATWTAGTAAGATGGACACCTGARACAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  |||||||||:|||||||| ||:|||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
               |||||:|||||||||||::||| ::||||||:||||||||:|||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:|||| |||||:||||| ||| ||||||||||||:: ||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:||||||||| ||||||||||||||||||||:||||::|||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||||| ||||||||||: |||||||||  |:|||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||:||||| |||||||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             ||||||||||||||:||||:|||||:||||||||||:|||| |||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 39
US-10-114-104-206
; Sequence 206, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
```

```
;              CITY: ALEXANDRIA
;              STATE: VA
;              COUNTRY: USA
;              ZIP: 22320
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/114,104
;              FILING DATE: 03-Apr-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/979,847
;              FILING DATE: 26-NOV-1997
;         ATTORNEY/AGENT INFORMATION:
;              NAME: BERRIDGE, WILLIAM P.
;              REGISTRATION NUMBER: 30,024
;              REFERENCE/DOCKET NUMBER: WPB 39046A
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-836-6400
;              TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 206:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1600 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: DNA (genomic)
;         SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-10-114-104-206

  Query Match             17.5%;  Score 1309.8;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||||| | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| ||||||| |||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||| | ||||||| || | |         ||| |  | | |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||||| |||  |||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |||||||:|||||||||||||:|||||:|:||:|||:::|||||:|||:||:|:|:::
Db        650 GYCARTYCTCATACCTGGACAYTCTTGTYCYTCAGTAYRKGGATGAYTTAMTTWTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||:|||||||||||||||:|| ||||||||||||:|||||||:|:|||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGSCAMCAAGCCACCCAAGHRCTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:::||||::|||||||||||||::||:|||:|||||  ||
Db        770 GTGGCTACAAGGTTTCCAAACMAARGGCTCRSCTCTGCTCACASSAGRTTARATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:||||||||||:|:|||:|||||||||:|||||||:::|||||||:|||||||||
Db        830 GGCTAAAATTATCCAAAGKCRCCAGGGCCCTCAGWGAGGAACGTATCCAGCSTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             :|||||| ||||:||:||||:|||||||||||||:::|||||||||:||::||:|||||
Db        890 MTTATCCMCATCCCAWAACCMTAAAGCAACTAAGARGGTTCCTTGGCATAWCAGSYTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||||::|||| |||:|:|:  ::::|||||||||::|||||::||||:||::||:|
Db        950 GCCGAAWAYRGATT-CCCVGRTACASYSMAATAGCCAGRCCATTATRTACAYTADYTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||||||:||:||:|||||||||:|:||||: :::|||||||||||
Db       1009 GAAACTCAGAAAGCCAATACCYATWTAGTAAGATGGACACCTGARACAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  |||||||:|||||||| ||:|||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||||  |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
              |||||:|||||||||||::||| ::||||||:|||||||:|||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:|||| |||||:||||| ||| ||||||||||||:: ||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:||||||||| ||||||||||||||||||:||||::|||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             |||||||||||||||||||| |||||||||||: |||||||||  |:||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| ||||||||||||||||||| || |:| ||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             |||||||||||||||||||:||:|||||||||||||||:||||| |||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             |||||||||||||||:||||:||||||:|||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600
```

```
RESULT 40
US-08-979-847-205
; Sequence 205, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  205:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1597 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-979-847-205

  Query Match             17.2%;  Score 1291.9;  DB 1;  Length 1597;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches 1405;  Conservative   88;  Mismatches  104;  Indels   19;  Gaps   10;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||| | |||||||||||||||| ||||||| || | |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| ||||||| |||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  | |||||||| || | |        |||    |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             | |||||||||| |||||||||||||| |||||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| |||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| ||| ||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |:||:|::|||||||||||||||||||||||:||| ||||||||||||||:| :|
Db        650 GTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCG 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||::|||||||||||||||||||||||||||||||| :|||||:| ||||||| |:||||
Db        710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:|:|||| :|||||||||||| || ||| ||||| ||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             ||||||||||||:|||||||||||||||||||||:|::|||||||||:|||:|||:||||
Db        890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||:|  |||| |||||||:  :|  |||||||||  :|||||:|||||:|||||| |
Db        950 GCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||| |||||:|||||||||:|:||    :::||||||||||||
Db       1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
             |||||||  ||  | | |||  ||||||||:||||||||| ||:||||||||:|||||||
Db       1066 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1123

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
             | ||||:||:|||:: ||||| |||||||||:||:||||||:||||||||||  |||:||
Db       1124 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1182

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
               |||||:|||||||||::||| ::|||||:|||||||:||||||||||||||||||||
Db       1183 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1242

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
             |||||:|:|||| ||||||:||||| ||| |||||||||||:: ||||||||||||||||
Db       1243 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1302

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
             |||||||||||||:||||||||| ||||||||||||||||||:|||||::|||||||||
Db       1303 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1362
```

```
Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
             ||||||||||||||||||||| |||||||||||: ||||||||| |:||||||||||||||
Db       1363 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1422

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
             ||||||: |||||||||| :|||| |||||||||||||||||| || |:| |||||||||
Db       1423 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1482

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
             ||||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||
Db       1483 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1542

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||
Db       1543 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1597


RESULT 41
US-10-114-104-205
; Sequence 205, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 205:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1597 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 205:
US-10-114-104-205

  Query Match             17.2%;  Score 1291.9;  DB 1;  Length 1597;
  Best Local Similarity   86.9%;  Pred. No. 0;

  Matches 1405;  Conservative   88;  Mismatches  104;  Indels   19;  Gaps   10;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||| |||||||||||||||||| || ||||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |||||||||||||||||||||||| | |||||||||||||||||| ||||||| || |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||||| |||||| ||||||| ||||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             |||||||||||  |  ||||||| || | |        |||   |  | |  |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||| |||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             ||||||||:|||||||||||:|||||||||| |||||||||||||||| ||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||| |||  ||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |:||:|::|||||||||||||||||||||||||:|||  |||||||||||||||:| :|
Db        650 GTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCG 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
             ||::|||||||||||||||||||||||||||||||| :|||||:| |||||||  |:||||
Db        710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

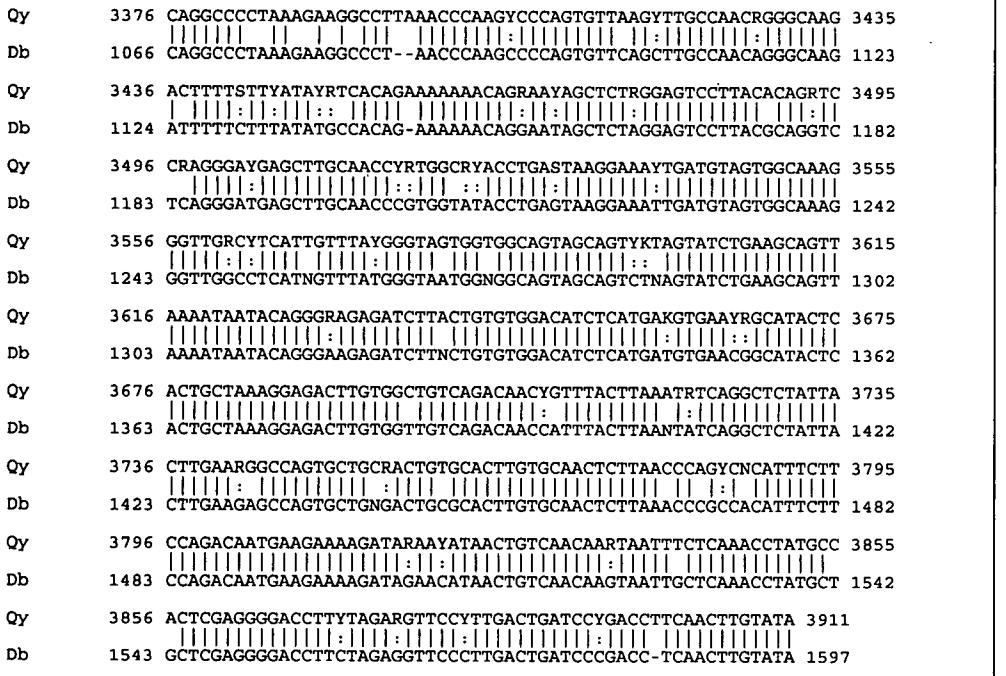
Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
             |||||||||:|||||||||||:|:|:|||| :|||||||||||| || ||| ||||| ||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
             ||||||||||||:|||||||||||||||||||||:|::|||||||||:|||:|||:||||
Db        890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

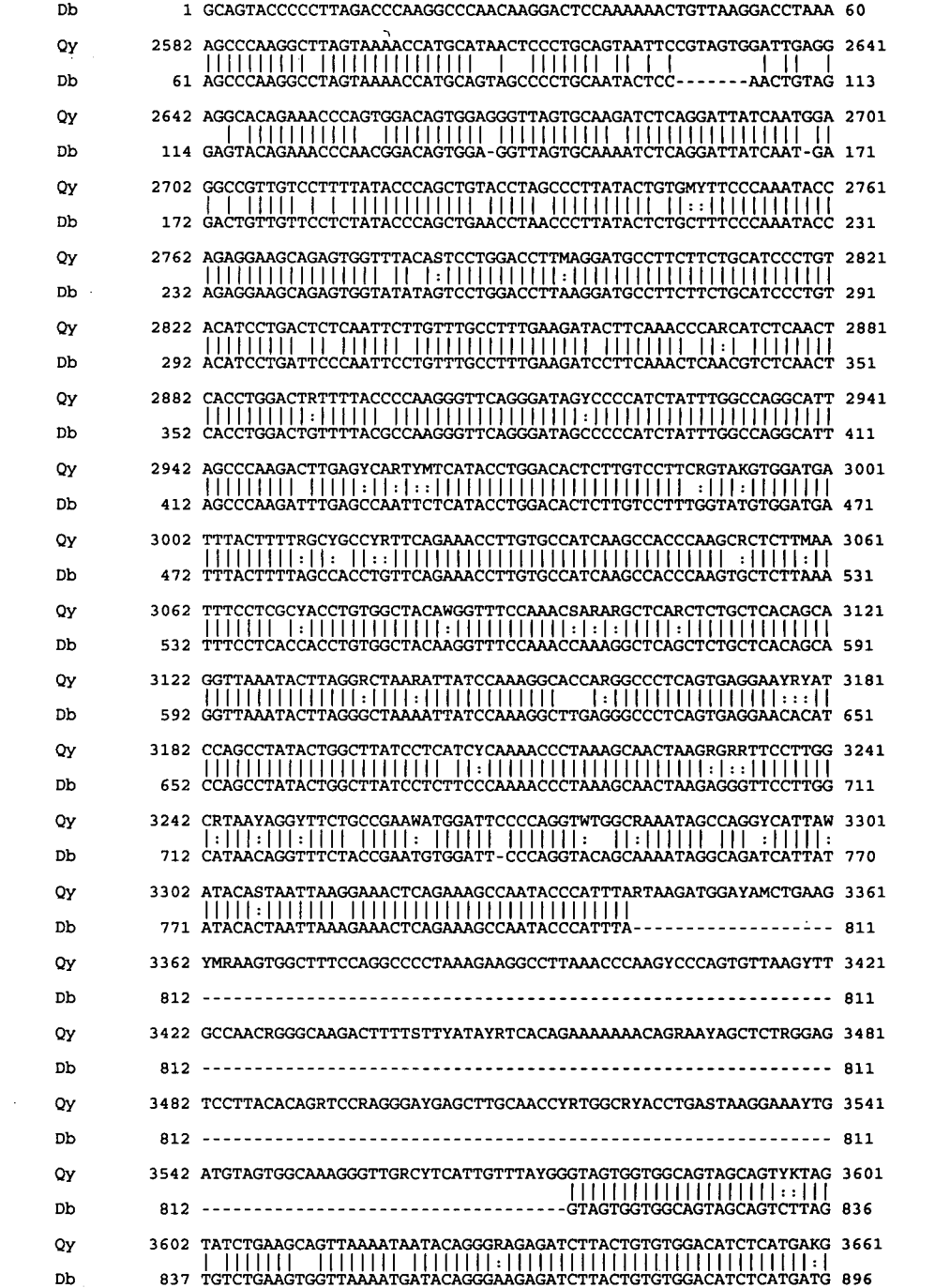
Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
             ||||||:|  |||| ||||||| :|  |||||||| :|||||:|||||:|||||| |
Db        950 GCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
             |||||||||||||||||||| |||||:|||||||||:|:||    :::|||||||||||||
Db        1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065
```



```
RESULT 42
US-10-450-763-9412
; Sequence 9412, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 9412
;   LENGTH: 1660
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2)..(994)
;   OTHER INFORMATION: 72% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1188.
US-10-450-763-9412

  Query Match             17.2%;  Score 1285.7;  DB 1;  Length 1660;
  Best Local Similarity   77.2%;  Pred. No. 0;
  Matches 1470;  Conservative   92;  Mismatches   98;  Indels  245;  Gaps    6;
```



```
Qy       3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
             ||||:: |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        897 TGAATGCCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAAT 956

Qy       3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
             :|||||||||||||||||||:|||||||||||||:|||| |||||||||| ||||||||||
Db        957 ATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACC 1016

Qy       3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTT 3841
             |||:| |||||||||||||||||||||||||||||:||:||||||| |||||||:|||||
Db       1017 CAGTCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTG 1076

Qy       3842 CTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTC 3901
             |||||||||||||| |||||||||| ||:||||:|||||:||||||||||: ||| ||
Db       1077 CTCAAACCTATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TC 1135

Qy       3902 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
             |||||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||
Db       1136 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCA 1195

Qy       3962 GTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAG 4021
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||:||
Db       1196 GTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAG 1255

Qy       4022 CTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAAT 4081
             ||:|||||||||||||||:|||:| ||| |||||||| ||||||||:||||||||:|||| 
Db       1256 CTGGCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAAT 1315

Qy       4082 ATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARA 4141
             ||||||||||||||||:|:| ||||:|||||||| |||||||||||||::|||||||||:|:| 
Db       1316 ATATATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGA 1375

Qy       4142 GAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARA 4201
             ||||||||||||||||||:|||:|||||||||||:|||||||||||||||||||||||:| 
Db       1376 GAAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAA 1435

Qy       4202 TTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCAN 4261
             ||||||:||||:|||||||||||||:||||||||:||| |||||||||:|||||||||| 
Db       1436 TTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAG 1495

Qy       4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
             |||||||||:|||||||||||:|||:|| :  |||||  :|||||||| :|||||||| 
Db       1496 AAAGGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCC 1555

Qy       4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
             |||||| |||||||||||||||||||||||||    |   ||| |||||| ||||||||||
Db       1556 ACAAGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCC 1615

Qy       4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
              |||||||| |||||||||||||||||||||||||| ||||| |
Db       1616 CTCCGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAG 1660
```

RESULT 43
US-10-450-763-12263
; Sequence 12263, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 12263
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(684)
; OTHER INFORMATION: 92% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93564,Smith-Waterman Score=1060.
US-10-450-763-12263

Query Match 17.2%; Score 1285.7; DB 1; Length 1660;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1470; Conservative 92; Mismatches 98; Indels 245; Gaps 6;

```
Qy       2522 GCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAA 2581
             |||||||||||| |||||||||||||||||||| ||||||| | |||||||||| ||||
Db          1 GCAGTACCCCCTTAGACCCAAGGCCCAACAAGGACTCCAAAAAACTGTTAAGGACCTAAA 60

Qy       2582 AGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGG 2641
             |||||||||| |||||||||||||| | ||||||| || | |         | ||  |
Db         61 AGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAATACTCC-------AACTGTAG 113

Qy       2642 AGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGA 2701
               | |||||||||||  |||||||||| |||||||||| |||||||||||||||||| ||
Db        114 GAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAAATCTCAGGATTATCAAT-GA 171

Qy       2702 GGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACC 2761
             | | ||||| | | |||||||||||| ||||| |||||||||| ||::||||||||||||
Db        172 GACTGTTGTTCCTCTATACCCAGCTGAACCTAACCCTTATACTCTGCTTTCCCAAATACC 231

Qy       2762 AGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGT 2821
             |||||||||||||||| || |:|||||||||:|||||||||||||||||||||||||||||
Db        232 AGAGGAAGCAGAGTGGTATATAGTCCTGGACCTTAAGGATGCCTTCTTCTGCATCCCTGT 291

Qy       2822 ACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACT 2881
             ||||||||| || |||||| |||||||||||||||| |||||||| ||:| ||||||||
Db        292 ACATCCTGATTCCCAATTCCTGTTTGCCTTTGAAGATCCTTCAAACTCAACGTCTCAACT 351

Qy       2882 CACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATT 2941
             ||||||||||:|||||| ||||||||||||||||||:|||||||||||||||||||||||
Db        352 CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATT 411

Qy       2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             |||||||||| |||||:||:|::|||||||||||||||||||||||||| :|||:|||||||
Db        412 AGCCCAAGATTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGA 471

Qy       3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
             |||||||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:||
Db        472 TTTACTTTTAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAA 531

Qy       3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
             ||||||| |:|||||||||||:|||||||||||:|:|:|||||:|||||||||||||||
Db        532 TTTCCTCACCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCA 591

Qy       3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
             |||||||||||||||:||||:|||||||||||||   |:|||||||||||||||:::||
Db        592 GGTTAAATACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACAT 651

Qy       3182 CCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGG 3241
             |||||||||||||||||||||||| ||:|||||||||||||||||||:|::|||||||||
Db        652 CCAGCCTATACTGGCTTATCCTCTTCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGG 711

Qy       3242 CRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAW 3301
             |:|||:|||:|||| ||||||: |||||| |||||||:  ||:|||||| ||| :|||||:
Db        712 CATAACAGGTTTCTACCGAATGTGGATT-CCCAGGTACAGCAAAATAGGCAGATCATTAT 770

Qy       3302 ATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAG 3361
             |||||:||||||| ||||||||||||||||||||||||||||||||
Db        771 ATACACTAATTAAAGAAACTCAGAAAGCCAATACCCATTTA---------------------- 811

Qy       3362 YMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTT 3421

Db        812 ------------------------------------------------------------ 811
```

```
Qy       3422 GCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAG 3481

Db        812 ------------------------------------------------------------ 811

Qy       3482 TCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTG 3541

Db        812 ------------------------------------------------------------ 811

Qy       3542 ATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAG 3601
                                                 ||||||||||||||||||||::|||
Db        812 ----------------------------------GTAGTGGTGGCAGTAGCAGTCTTAG 836

Qy       3602 TATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKG 3661
                |  |||||||  ||||||||  ||||||||:||||||||||||||||||||||||:|
Db        837 TGTCTGAAGTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATG 896

Qy       3662 TGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAAT 3721
                ||||:: |||||||||||||||||||||||||||||||||||||:||||||||||||
Db        897 TGAATGCCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAAT 956

Qy       3722 RTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACC 3781
                :|||||||||||||||||||:|||||||||||:|||| |||||||| ||||||||||
Db        957 ATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACC 1016

Qy       3782 CAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTT 3841
                |||:| |||||||||||||||||||||||||||||:||:|||||| ||||||:|||||
Db       1017 CAGTCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTG 1076

Qy       3842 CTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTC 3901
                |||||||||||||| |||||||||| ||:||||:|||||:|||||||||||: ||| ||
Db       1077 CTCAAACCTATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TC 1135

Qy       3902 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
                |||||||||||||||||||||||||||||||||||||||||||||| |||||:||||||
Db       1136 AACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCA 1195

Qy       3962 GTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAG 4021
                |||||||||||||||||||:|||||||||||||||||||||||||||||||||||:||
Db       1196 GTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAG 1255

Qy       4022 CTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAAT 4081
                ||:|||||||||||||:|||:| ||| |||||||| ||||||||:||||||||:||||
Db       1256 CTGGCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAAT 1315

Qy       4082 ATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARA 4141
                |||||||||||||||:|:| ||||:|||||||| |||||||||||::|||||||:|:|
Db       1316 ATATATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGA 1375

Qy       4142 GAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARA 4201
                |||||||||||||||||||:|||:||||||||||||:|||||||||||||||||||:|
Db       1376 GAAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAA 1435

Qy       4202 TTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCAN 4261
                ||||||:||||:|||||||||||||:||||||||:||| |||||||||:|||||||||
Db       1436 TTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAG 1495

Qy       4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
                |||||||||:|||||||||||:|||:|| :  |||||  :|||||||| :|||||||||
Db       1496 AAAGGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCC 1555

Qy       4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
                ||||||  |||||||||||||||||||||||||||    |   |||  |||||| ||||||||||
Db       1556 ACAAGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCC 1615

Qy       4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
                ||||||||  |||||||||||||||||||||||||||||||| ||||| |
Db       1616 CTCCGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAG 1660


RESULT 44
US-10-450-763-20005
; Sequence 20005, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 20005
;   LENGTH: 1660
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2)..(994)
;   OTHER INFORMATION: 72% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1188.
US-10-450-763-20005

  Query Match             17.2%;  Score 1285.7;  DB 1;  Length 1660;
  Best Local Similarity   77.2%;  Pred. No. 0;
  Matches 1470;  Conservative   92;  Mismatches   98;  Indels  245;  Gaps    6;

Qy       2522 GCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAA 2581
                ||||||||||||  |||||||||||||||||||||||| |||||||  | ||||||||||  ||||
Db          1 GCAGTACCCCCTTAGACCCAAGGCCCAACAAGGACTCCAAAAAACTGTTAAGGACCTAAA 60

Qy       2582 AGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGG 2641
                ||||||||||  ||||||||||||||||  |  |||||||  || | |          | ||  |
Db         61 AGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAATACTCC-------AACTGTAG 113

Qy       2642 AGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGA 2701
                  | ||||||||||||  |||||||||| |||||||||||| |||||||||||||||||| ||
Db        114 GAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAAATCTCAGGATTATCAAT-GA 171

Qy       2702 GGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACC 2761
                | | ||||| | | |||||||||||| ||||| |||||||||| ||::|||||||||||||
Db        172 GACTGTTGTTCCTCTATACCCAGCTGAACCTAACCCTTATACTCTGCTTTCCCAAATACC 231

Qy       2762 AGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGT 2821
                |||||||||||||||| || |:|||||||||||:|||||||||||||||||||||||||
Db        232 AGAGGAAGCAGAGTGGTATATAGTCCTGGACCTTAAGGATGCCTTCTTCTGCATCCCTGT 291

Qy       2822 ACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACT 2881
                ||||||||| || |||||| |||||||||||||||| |||||||| ||:| ||||||||
Db        292 ACATCCTGATTCCCAATTCCTGTTTGCCTTTGAAGATCCTTCAAACTCAACGTCTCAACT 351

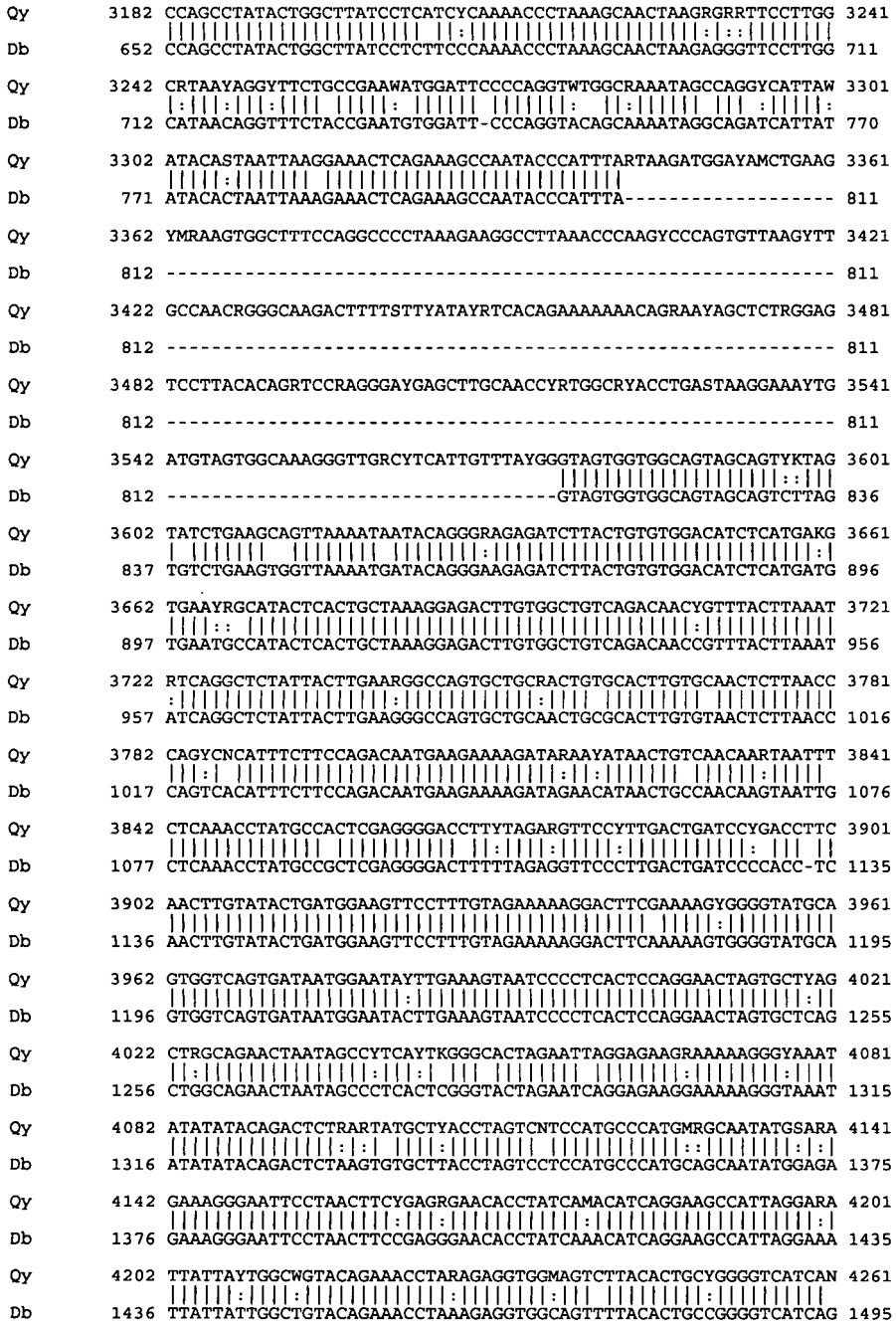
Qy       2882 CACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATT 2941
                ||||||||||:|||||| ||||||||||||||||||:|||||||||||||||||||||||
Db        352 CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATT 411

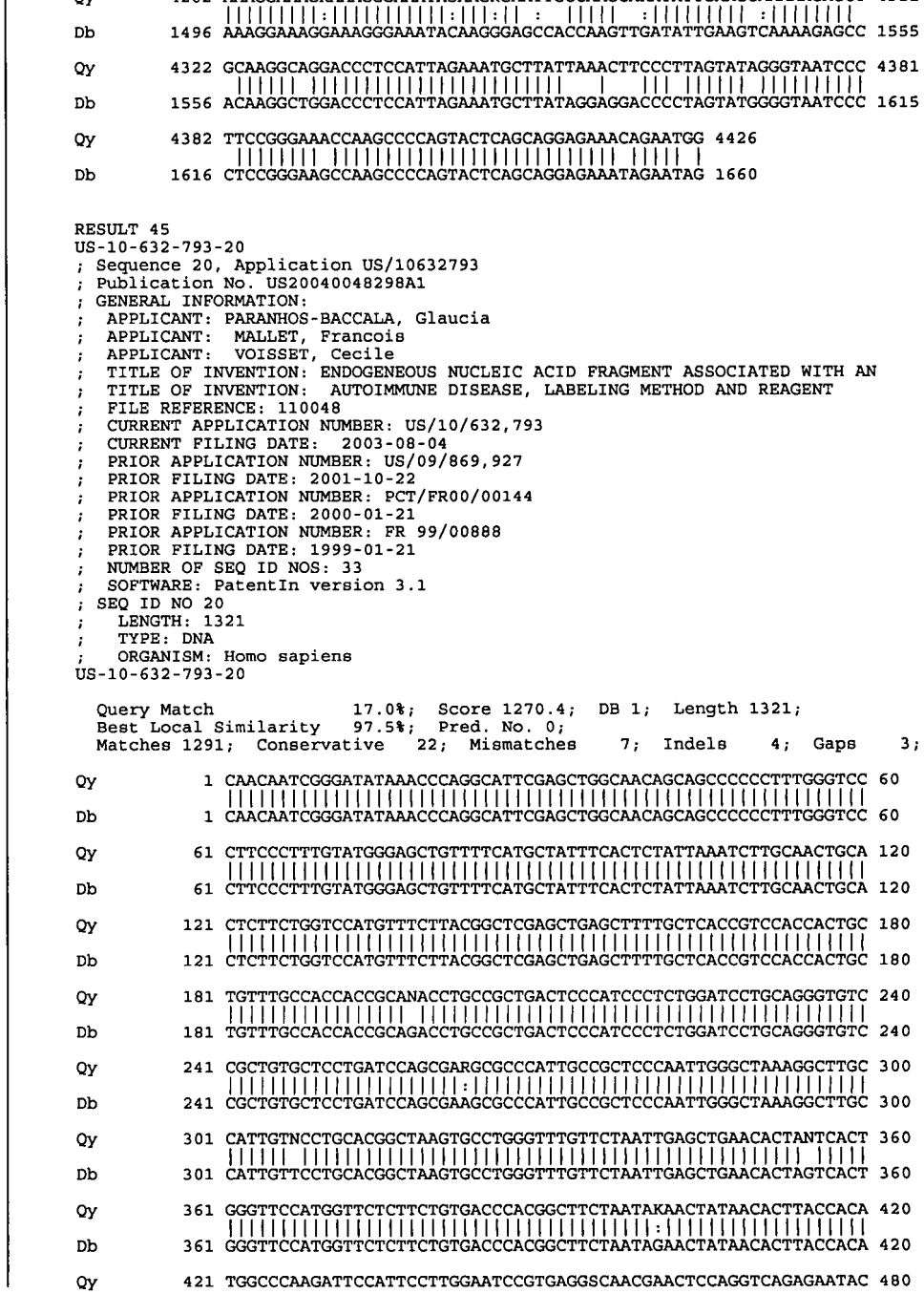
Qy       2942 AGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
                |||||||| |||||:||:|::|||||||||||||||||||||| :|||:||||||||
Db        412 AGCCCAAGATTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTATGTGGATGA 471

Qy       3002 TTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAA 3061
                |||||||||:||: ||::||||||||||||||||||||||||||||||| :|||||:||
Db        472 TTTACTTTTAGCCACCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAA 531

Qy       3062 TTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCA 3121
                ||||||| |:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||
Db        532 TTTCCTCACCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCA 591

Qy       3122 GGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT 3181
                |||||||||||||||:||||:|||||||||||||   |:|||||||||||||||:::||
Db        592 GGTTAAATACTTAGGGCTAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACAT 651
```





RESULT 45
US-10-632-793-20
; Sequence 20, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-20

Query Match 17.0%; Score 1270.4; DB 1; Length 1321;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1291; Conservative 22; Mismatches 7; Indels 4; Gaps 3;

```
                |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              |||||||||||||||||||||||||||||||:|||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:|||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              |||| |
Db       1317 GAGG 1320


RESULT 46
US-10-717-580-1
; Sequence 1, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1321
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-1

  Query Match            17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| |||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              |||||||||||||||||||||||||||||||:|||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719
```

```
Qy       721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy       781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy       841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy       901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
            ||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db       900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy       961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy      1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db      1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy      1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
            |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy      1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
            |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db      1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy      1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy      1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
            ||||||||:|||||||||||||||||||||||||||||||||     ||||||||||
Db      1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy      1321 GAAG 1324
            || |
Db      1317 GAGG 1320


RESULT 47
US-09-864-761-14940
; Sequence 14940, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14940
;   LENGTH: 1472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007567.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
US-09-864-761-14940

  Query Match             16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy      3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
            :|||||:|||||||||| | |||| ||||||||||||:||||:|||||||||||||||||:
Db         1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy      3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
            ||||||||||||||||:::|||||||| ||||||||||||||||||:||||||||||||
Db        61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy      3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
            |||||||||:|::||| |||||:|| :|||:|||||||||:|||||||| |||||||: |
Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy      3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
            ||:|||||| |||:|||||:|||||:||||| ||||||||||||||||||||||||||||
Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy      3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
            ||:|||||||||:|:||||||:::||| |||||||||||||  ||  | | |||  |||
Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy      3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
            |||||:|||||||    ||||:||||||||:|||||||||||| :||:|||::||||||
Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy      3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
            |||||||||:||:||||||:|||||||||||||||:|||:|||||: ||| |||||||::
Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy      3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
            ||||::||||||:||||||||:|||||||||||||||||||:|:||||||||: ||
Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476
```

```
Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
             |||||| ||||||||||||::|||||||||||||||||||| ||||||||| :|||||||
Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
             ||||||||||||||||||||:|||||::|||||||||||||||||||||||||||||||||
Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
             |||||||||||:||||||||||||:|||||||||| | ||||||:| ||||||||||:||
Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
             |||||||||||||||||||||||:| |||||||||||| ||||||||||||||||||||
Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
             ||:|||||||||||||:|||||  |||||||| |||||| |||||||||||:||||:|
Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
             ||||: |||||||||:|||| ||||||||||||||||||||||||| |||||||||
Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
             || |||| |||||:|||||||||||| |||||||||||||||:|||||||||||||||||
Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
             |||||||||||||||||||:||||:|||||||||||||||:|||:|  |||| |||||
Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
             ||   |||||:||||||||:|||||||||||||||||||:|:||||||:|||||||| ||
Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
             ||||||||||::|||||||||:|:|||||||| |||||||||||:|||:|||||||||||
Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
             |:|||||||||||||||||||:|||||||:|||| ||||||||||||:|||||||||:|
Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
             ||||||||||||:|||||||||| |||||||||:||||||||||:|||:|||:||||||
Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
             || :| |||||||||:|||||||||||||||||||||||||||||||||||||||||  ||
Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
                 ||| ||||||  |||||||||| || |||||||||||||||||||||  || ||||
Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
             | |||||||||||||||| ||||    ||||| |||||||| ||| || |||||||  ||
Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
             |||||  |||||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
             ||||||||||||||||  |||| ||||||| |:||| |||
Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472


RESULT 48
US-11-028-539-2
; Sequence 2, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/11/028,539
;  CURRENT FILING DATE:  2005-01-05
;  PRIOR APPLICATION NUMBER: US/09/719,554
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1326
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-028-539-2

  Query Match            16.0%;  Score 1201.4;  DB 1;  Length 1326;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1295;  Conservative   16;  Mismatches   14;  Indels   12;  Gaps    9;

Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCCTTCAACCCAAA 957
              ||||||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
              |:||||::|||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
              ||||||||||:|||||||||||||||||||||||||||||||||     |||||||||
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
              |||| |||||||| |||||||||||||||||||:|||||||||||||:|||||||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
              |||||   ||| | |     ||||| ||||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711
```

```
Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
             |||||:|||||||||||||||||:|||| ||:||||:|||||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
             |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009

Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
             |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
             |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326


RESULT 49
US-08-979-847-117
; Sequence 117, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  117:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-117

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| |||||||  ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| |||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||| || |||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| |||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||||| ||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| ||||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |  ||||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716
```

```
Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
             |||| | ||  | ||      |||||||||||||||:||||||||||| ||:||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
             ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
             ||||:||||||| ||||||||||:|||| ||: |||:||||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
             ||||||| ||||| |||||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
             ||||||| | |||| |||||||||||||||  |||||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
             |||| |||||||||||||||| |||||||| |:|| ||||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
             |||||||||||| ||| |||||||||||||| ||||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
             ||||  ||||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
             |||||||| |||||||||| | |||||||||||||| ||||||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
             ||||||||||||||||  ||| |  | |||||||||||||| ||||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
             |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
             |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 50
US-10-114-104-117
; Sequence 117, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 117:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1511 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-10-114-104-117

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||:|||| |||| ||||||| ||||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               ||||||||||||||||||||||||| ||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || |||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||      | |||||||||||||| ||::|||||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
```

```
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| ||||||||||:|||||||||||||||||||||||| |||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||  |  |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      |||||||||||||:||||||||||| ||:||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| ||||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| |||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| ||||||||||||||||  |||||||||||||||||||||| ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||||  ||| |||||||||||||  ||||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| ||||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | ||||||||||||||  |||||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| |||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

```
RESULT 51
US-10-632-793-1
; Sequence 1, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-1

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches   118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||||||| |||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||||| || |||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||      | |||||||||||||||| ||::|||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| ||||||||||:|||||||||||||||||||||||| |||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599
```

```
Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
             ||||| |||| ||| ||||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
             |||||| |||:|||||||||||| |||||||||||||||| |||    |   ||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
             |||| |  ||  | ||       |||||||||||||:||||||||||| ||:||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
             ||||   ||| | |     ||||| ||||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
             ||||:||||||| ||||||||||:|||| ||: |||:|||||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
             ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
             ||||||| | |||| |||||||||||||| ||||||||||||||||||||   ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
             |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
             |||||||||| ||| |||||||||||||| ||||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
             |||| ||||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
             |||||||| |||||||||| | |||||||||||||||  |||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
             ||||||||||||||||  ||| |  | |||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
             |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
             |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

```
RESULT 52
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-25

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             |||||||||||||||:|||||||||||||||||||||:|||||||||||:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:|||||||||:||||||||||||||||||:||||||:|||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 53
US-10-717-580-6
; Sequence 6, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-6

  Query Match            14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             ||||||||||||||:|||||||||||||||||||:||||||||||||:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:||||||||:||||||||||||||||:||||||:||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 54
US-11-028-539-28
; Sequence 28, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
```

```
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/11/028,539
;  CURRENT FILING DATE:  2005-01-05
;  PRIOR APPLICATION NUMBER: US/09/719,554
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
;   LENGTH: 1080
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1080)
;   OTHER INFORMATION:
US-11-028-539-28

  Query Match             12.8%;  Score 960.2;  DB 1;  Length 1080;
  Best Local Similarity   96.4%;  Pred. No. 0;
  Matches 1052;  Conservative   16;  Mismatches   11;  Indels   12;  Gaps    9;

Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
              ||||||||||||||||||||||||||||||||||||||::||||||||||||||||||||
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
              ||||||||:||||::|||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
              ||||||:||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||     |
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
              |||||||||| |||||||| ||||||||||||||||||:||||||||||||||||||:
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
              ||||||||||||   ||| | |      ||||| ||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
              ||||||||||||:||||||||||||||||:|||| ||:||||:|||||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db         890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
              |||||||||||
Db        1070 TGAATGCCTAA 1080


RESULT 55
US-08-979-847-57
; Sequence 57, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1167 base pairs
```

```
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-57

  Query Match            12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
              ||||||||||||||||||||||||||||||:|||||||||:|:| ||||:::|||  |||
Db             1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
              ||||||||||  ||  ||  |||  ||||||||:|||||||||||:||||||||:||||
Db            61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
              |||||||||:||:|||::||||||  ||||||||:||:|||||||:||||||||||||||
Db           119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
              :|||:|||||: ||||||||||||::||||::|||||||:||||||||:||||||| |||||
Db           177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
              ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| ||||||  ||
Db           237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
              ||||||||||||||||:|||||||||||||||||||||||||||||:||||||::|||||
Db           296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
              |||||||||||||  |||||||||||||||||||||: ||||||||||: |||| ||||
Db           356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
              |||||||||: ||||||||||:|||||  || ||||||||||||||||||||:| |||||
Db           416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
              |||||||||||||||||||||||||:||:|||||||||||||||:|||||  ||||||||||
Db           476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
              ||  |||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||||
Db           536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
              |||||||||||||| | |||||||||||||  ||||||:|||||||||||||| |||||||
Db           595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
              ||||||||||:|||||||||||||| |||||||||||||||||||||:| ||:|||||||
Db           655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
              |||||||:|||:|:|||||||||||||||||||:||||||||:|||||||||| ||||
Db           715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
              ||||:|:||||||:|||||||| ||||||||||||::||||||||:|:||| ||||||||
Db           775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
              |||||||||:|||:|||||||||||||: ||||| |||||||||||||:|||||||:|||
Db           835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
              |:|||||||||||:||||||||:|||||||||||||: |||||||||  ||| | ||:
Db           895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
              ||||||||||:|||: ||: ||||||| :|||||||||:|||| ||| ||||||||||
Db           955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
              || ||||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db           1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
              |||||||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db           1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
              |||||||  || || |||||||||| ||||| |
Db           1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 56
US-10-430-442-61
; Sequence 61, Application US/10430442
; Publication No. US20030186391A1
;   GENERAL INFORMATION:
;        APPLICANT: Herve PERRON
;                   Frederic BESEME
;                   Frederic BEDIN
;                   Glaucia PARANHOS-BACCALA
;                   Florence KOMURIAN-PRADEL
;                   Colette JOLIVET
;                   Bernard MANDRAND
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 92
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Oliff & Berridge
;             STREET: 700 South Washington Street, Suite 300
;             CITY: Alexandria
;             STATE: Virginia
;             COUNTRY: U.S.A.
;             ZIP: 22314
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/430,442
;             FILING DATE: 07-May-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/691,563
;             FILING DATE: 02-AUG-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Berridge, William P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 38588
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 61:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1167 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-10-430-442-61

  Query Match            12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;
```

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db        119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db        177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db        356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db        416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db        476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db        536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db        595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167
```

RESULT 57
US-10-114-104-57
; Sequence 57, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,847
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-10-114-104-57

Query Match 12.7%; Score 952.8; DB 1; Length 1167;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1013; Conservative 77; Mismatches 73; Indels 10; Gaps 6;

```
Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118
```

```
Qy      3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
             |||||||||:||:|||::||||||  ||||||||:||:||||||:||||||||||||||||
Db       119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy      3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
             :|||:|||||: |||||||||||::||||::||||||:||||||||:||||||| |||||
Db       177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy      3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
             ||||||||:|:||||||||||: ||||| ||  ||||||||||::|||| |||||| ||
Db       237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy      3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
             ||||||||||||||||:||||||||||||||||||||||||||||:|||||::|||||
Db       296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy      3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
             |||||||||||||  |||||||||||||||||||||: ||||||||||: |||| ||||
Db       356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy      3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
             |||||||||: ||||||||||||:||||  || |||||||||||||||||||:| |||||
Db       416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy      3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
             |||||||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||
Db       476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy      3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
             ||  ||||||||||||||:||||:|||||:||||||||||||:|||| ||||||||||||
Db       536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy      3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||| | |||||||||||||| ||||||:|||||||||||||| |||||||
Db       595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy      3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||| ||||||||||||||||||||:| ||:||||||||
Db       655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy      4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:|:|||||||||||||||||||:|||||||:|||||||||| ||||
Db       715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy      4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:||||||:|||||||| ||||||||||||::|||||||:|:||| ||||||||
Db       775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy      4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||:|||:||||||||||||: ||||| |||||||||||||||:|||||||:|||
Db       835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy      4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:|||||||:|||||||||||||||: |||||||||  ||| | ||:
Db       895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy      4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||: ||: ||||||| :||||||||||:|||| ||| ||||||||||
Db       955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy      4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             || |||||||||||||||||  ||   ||| |||||| |||||||||| || ||||||
Db      1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy      4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| |
Db      1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy      4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
             |||||||  || || |||||||||| ||||| |
Db      1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167
```

```
RESULT 58
US-09-864-761-31445
; Sequence 31445, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31445
;   LENGTH: 1117
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007567.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P26808, EVALUE 4.00e-64
US-09-864-761-31445

  Query Match           12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;
```

```
Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:||||||||:|:|||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             |||||||| ||  | | |||  ||||||||:||||||||    ||||:||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             ||||||||| :||:|||::|||||| |||||||||:||:||||||:|||||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::|||||::|||||:|||||||:||||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:||||||||||: |||||||| ||||||||||||::|||||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| |||||||| :|||||||||||||||||||||||||||||:||||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | ||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             |||||||  ||||||||||:||||:|||||: ||||||||||:|||| ||||||||||||
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:|||||||||||||| |||||||
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             ||||||||||:|||||||||||||||||||||||||||||||||||:||||:|||||||
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   |||||:|||||||:|||||||||||||||
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:|||||||:|||||||| ||||||||||::|||||||:|:|||||||| ||
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:|||||||||||||:|||||||||||||||||||||:|||||||:|||
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | ||||||||||||||:|||||||||||||||||||:|||||||||| ||||||||||:
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|||:|||||||| :| ||||||||:||||||||||||||||||||
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||| ||   ||| |||||| |||||||| || ||||||
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
                  |||||||||||||||| || ||||| ||||||||||||||||
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117
```

RESULT 59
US-10-029-386-23846
; Sequence 23846, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23846
; LENGTH: 1117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007567.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P26808, EVALUE 4.00e-64
US-10-029-386-23846

Query Match 12.7%; Score 948; DB 1; Length 1117;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 988; Conservative 74; Mismatches 52; Indels 10; Gaps 5;

```
Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:||||||||:|:|||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             |||||||| ||  | | |||  ||||||||:||||||||    ||||:||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             ||||||||| :||:|||::|||||| |||||||||:||:||||||:|||||||||||||||
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::|||||::|||||:|||||||:||||||||||||||
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:||||||||||: |||||||| ||||||||||||::|||||||||||||||
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| |||||||| :|||||||||||||||||||||||||||||:||||||::||||
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             ||||||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | ||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537
```

```
Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 60
US-08-979-847-1
; Sequence 1, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790
```

```
Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| |||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 61
US-10-430-442-1
; Sequence 1, Application US/10430442
; Publication No. US20030186391A1
;   GENERAL INFORMATION:
;        APPLICANT: Herve PERRON
;                   Frederic BESEME
;                   Frederic BEDIN
;                   Glaucia PARANHOS-BACCALA
;                   Florence KOMURIAN-PRADEL
;                   Colette JOLIVET
;                   Bernard MANDRAND
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 92
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Oliff & Berridge
;             STREET: 700 South Washington Street, Suite 300
;             CITY: Alexandria
;             STATE: Virginia
;             COUNTRY: U.S.A.
;             ZIP: 22314
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/430,442
;             FILING DATE: 07-May-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/691,563
;             FILING DATE: 02-AUG-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Berridge, William P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 38588
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1158 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-430-442-1

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  |||||||| |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:|||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:||||||||||::||| ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
```

```
                |||||:|||||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:||||||||| |||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                |||||  |:||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                ||||| ||
Db       1150 CTTAAACC 1157


RESULT 62
US-10-114-104-1
; Sequence 1, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1158 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-114-104-1
```

```
  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  ||||||||  ||||||||||  ||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                ||||  |||||  |||||  | | ||||||||||||||||||  |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||||  |||||||||||||||||||||||||||||||||||  |||  |||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| |||||||  |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                |||||  ||  |||  |||||  |||:||||:|||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                |||||||||:|||:|||:|||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :||||:|||||:|||||| |||||||||||||||||||||||| |||||:||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:|||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
                | ||:|||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:|||||||||||  |||:||  |||||:|||||||||||::|||  ::||||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:|||||||||| |||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029
```

```
Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:||||||::||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:||||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 63
US-11-028-539-7/c
; Sequence 7, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/11/028,539
;  CURRENT FILING DATE:  2005-01-05
;  PRIOR APPLICATION NUMBER: US/09/719,554
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 1216
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-028-539-7

  Query Match             11.9%;  Score 894.401;  DB 1;  Length 1216;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches 1003;  Conservative   21;  Mismatches  102;  Indels   60;  Gaps    3;

Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATT 4729
             |||||||||||||:||||||:|||||||||||||| | :|:||:|||||||||| ||||
Db       1216 AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACTCAGGAGAAGACTGGCAACTAGATT 1157

Qy       4730 TTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCA 4789
             |||||||| ||||||| ||||||||||||||||||||||||||||||:||| |||||
Db       1156 TTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATATTTTCA 1097

Qy       4790 CGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCAC 4849
             | |||||||||:|||||||||  ||||||||||||||| |||||| ||||||||||||||
Db       1096 CTGGTTGGGCAGAGGCCTTCCTTTGTAGGACAGAAAAGGACCAAGATGTAATAAAGGCAC 1037

Qy       4850 TAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCC 4909
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1036 TAATTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATGGCC 977

Qy       4910 CTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACT 4969
             | | |||| |||| |||||||||||||||||||||||||| ||||| |||| ||||||||
Db        976 CCGTTTTCAAGGCTACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCACT 917

Qy       4970 TACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCA 5029
             |||||||||| || |||||||| |||||||| || ||| |||||||||||||| :|||||
Db        916 TACACTGCGCGTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAATGAAGCACTCA 857

Qy       5030 AAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTA 5089
             || ||||| |||||||| |||||| |||||||||| |||||||||:||||||||||
Db        856 AACGACATTTAAAAAAGCTAACCCAAGAAACCCACCTTGCATGGCCTGCTCTGTTGCCTA 797

Qy       5090 TAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGC 5149
              |||||||  |||||||  |||| ||||||||||||||||||||||||| ||||||
Db        796 CAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAAGCAGGACTTAGCCCATACAAAATGC 737
```

```
Qy       5150 TGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAG 5208
             |||||||| |||||||| ||||||| |||||||||||||||| | |||  |||| |||||
Db        736 TGTATGGATGGCCCTTCCTAACCAAGGACCTTGTGCTTGACCGAGAGATGGCCAGCTTAG 677

Qy       5209 TTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCT 5268
             |||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||||
Db        676 TTGCAGACATCACCTCCTTGGCCAAATATCAACAAGTTCTTAAAACATTACAGGGAACCT 617

Qy       5269 ATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCC 5328
             | || |||||||||||||  || |||| |||||: |||||||||||||||||||||||||
Db        616 GTTCCCGAGAAGAGGGAAAGTAATTATTGCACCCTGGTGACATGGTATTAGTCAAGTCCC 557

Qy       5329 TTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 5388
             |||: ||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db        556 TTCCTTCTAAGTCCCCATCCTTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 497

Qy       5389 YTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 5448
             :|||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        496 CTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGAATACATCACACTCGAGTCA 437

Qy       5449 AATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 5508
             ||  ||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        436 AACTCTGTATACTGCCAAAGGAACCCGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 377

Qy       5509 AACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAAT 5568
             |||||||||||||| ||| |||||||||||| | |||||  |||||||||||||| |||
Db        376 AACCTCTAGAGGATCTGCACCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAAT 317

Qy       5569 CATAAATCCCCCATG---------------------------------------------- 5583
             | || |||| |   |
Db        316 CGTAGATCCTCATGGCCCTCCTTGTCATATTTTTTTCTTTACTGTTCTCTTACCCCCTTT 257

Qy       5584 -------------GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTT 5630
                         :|||||||||||||||||||||||:||:||||:|||:||||||:|||
Db        256 CACTCTCACTGAACCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTT 197

Qy       5631 CACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTT 5690
             |||||||||||||||||||||||||| |||||  |||||||||||| |:|||:|||| |
Db        196 CACTCTCACTGCACCCCCTCCATGCCACTGTACTACCAGTAGCTCCCATTACCAAGAGCT 137

Qy       5691 TCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAA 5750
             |||||||||||||| || ||| |||||||||||||| |||||||||||||||| |||:|||
Db        136 TCTATGGAGAATGCGGCTTCCTGGAAATATTGATGACCCATCGTATAGGAGTTTTTCTAA 77

Qy       5751 GGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||| |||||  |||||  |||||||| |||||  |||   ||| |
Db         76 AGGAAACCCCATTTTCACCACCCACACCTATATGACCCTGCACTTC 31


RESULT 64
US-10-301-480-534556/c
; Sequence 534556, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 534556
;   LENGTH: 990
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-534556
```

Query Match 10.4%; Score 777.2; DB 1; Length 990;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 864; Conservative 42; Mismatches 76; Indels 14; Gaps 6;

```
Qy       3994 TCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACT 4053
             |||||||||||||||||||| |||:||||:|||||||||||||||:|||:|  ||||||
Db        990 TCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCAGGCACT 931

Qy       4054 AGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGT 4113
             ||||||||| || |:||||| ||:|||||||||||||| |||:|:||||||:|||||||
Db        930 AGAATTAGGGGACGGAAAAATGGTAAATATATATACAGAATCTAAGTATGCTTACCTAGT 871

Qy       4114 CNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACC 4173
             | |||| |||||||::|||||||:|:|||||||||||||||||||:|||:|||||||
Db        870 CCTCCACGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGGAACACC 811

Qy       4174 TATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGT 4233
             |||||:|||||||||||||||||:|||||||:||||: || |||||||||:||||||
Db        810 TATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTATAGAGAAACCTAAAGAGGT 751

Qy       4234 GGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTG 4293
             ||:|||||||||||||:|||||||||     ||||||:||||||||||:|||:|||: |
Db        750 GGCAGTCTTACACTGCTGGGGTCATC-----GGAAAGGAAAGGGAAATAGAAGGGAACCG 696

Qy       4294 CCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTAT 4353
             |||||  :||||||||||:|||||||||  ||||||||||| ||||||||||||||||||
Db        695 CCAAGTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACACTCCATTAGAAATCCTTAT 636

Qy       4354 TAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGG 4413
               ||   ||| |||||| |||||||||| || ||||||||||||||||||||| |||||
Db        635 AGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCTGCAGG 576

Qy       4414 AGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCAC 4470
             ||| | |||||||||||||||| |||    ||||| ||| |||| ||| || |||||||
Db        575 AGACATAGAATGGGGAACCTCACAAGGACATAGTTTCCTCTCCTCAGGATGGCTAGCCAC 516

Qy       4471 TGAAGAAGGGAAA---ATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCT 4527
             |||||||||  ||    |  ||   ||||||| ||| ||||| ||||||||||||||||||
Db        515 TGAAGAAGGACAACAGAACCTAAAGCCTGCAGCTAACCAATAGAAATTACTTAAAACCCT 456

Qy       4528 TCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTAC 4587
             ||| |||||||||||||||||| ||||||||||||||:|||||||||||||||||||||
Db        455 TCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTAC 396

Qy       4588 TGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARAR 4647
             |||||||||||||||||||||||||||||:|||:|||||||||||||:|||||||:|:
Db        395 TGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAAAG 336

Qy       4648 AAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCC 4707
             ||||||||||||||:||| |||||||||||||:|| |||:|||||||||||||||||
Db        335 AAATAATCCCCTGCCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCC 276

Qy       4708 TGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCT 4766
              |:|:||:|||||||||| |||||||||||| ||||||| |||| ||||||||||||||||
Db        275 AGGAGAAGACTGGCAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCT 217

Qy       4767 ACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAA 4826
             ||||||  || ||:||||||| || ||||||| |:||||||||| ||| |||| ||||||
Db        216 ACTAGTTCGGCTAGATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAA 157

Qy       4827 GGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCG 4886
             ||||| |||| |||||||| ||||||||||||||||||||||||||||||||||| |||| 
Db        156 GGCCCGAGAGATAATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCG 98

Qy       4887 AGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCA 4946
             |||||||||||||||||| |||| |||||| |||| ||||||||||||||||||||||||
Db         97 AGGCTTACAGAGTGACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCA 38

Qy       4947 GGCGTTAGGTATACGATATCACTTACACTGCGCCTG 4982
             ||| ||||| |||| |||||||||||| |||||||
Db         37 GGCATTAGGCATACAATATCACTTACACCGCGCCTG 2
```

RESULT 65
US-10-301-480-1147965/c
; Sequence 1147965, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1147965
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1147965

Query Match 10.4%; Score 777.2; DB 1; Length 990;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 864; Conservative 42; Mismatches 76; Indels 14; Gaps 6;

```
Qy       3994 TCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACT 4053
             |||||||||||||||||||| |||:||||:|||||||||||||||:|||:|  ||||||
Db        990 TCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCAGGCACT 931

Qy       4054 AGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGT 4113
             ||||||||| || |:||||| ||:|||||||||||||| |||:|:||||||:|||||||
Db        930 AGAATTAGGGGACGGAAAAATGGTAAATATATATACAGAATCTAAGTATGCTTACCTAGT 871

Qy       4114 CNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACC 4173
             | |||| |||||||::|||||||:|:|||||||||||||||||||:|||:|||||||
Db        870 CCTCCACGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGGAACACC 811

Qy       4174 TATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGT 4233
             |||||:|||||||||||||||||:|||||||:||||: || |||||||||:||||||
Db        810 TATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTATAGAGAAACCTAAAGAGGT 751

Qy       4234 GGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTG 4293
             ||:|||||||||||||:|||||||||     ||||||:||||||||||:|||:|||: |
Db        750 GGCAGTCTTACACTGCTGGGGTCATC-----GGAAAGGAAAGGGAAATAGAAGGGAACCG 696

Qy       4294 CCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTAT 4353
             |||||  :||||||||||:|||||||||  ||||||||||| ||||||||||||||||||
Db        695 CCAAGTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACACTCCATTAGAAATCCTTAT 636

Qy       4354 TAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGG 4413
               ||   ||| |||||| |||||||||| || ||||||||||||||||||||| |||||
Db        635 AGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCTGCAGG 576

Qy       4414 AGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCAC 4470
             ||| | |||||||||||||||| |||    ||||| ||| |||| ||| || |||||||
Db        575 AGACATAGAATGGGGAACCTCACAAGGACATAGTTTCCTCTCCTCAGGATGGCTAGCCAC 516

Qy       4471 TGAAGAAGGGAAA---ATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCT 4527
             |||||||||  ||    |  ||   ||||||| ||| ||||| ||||||||||||||||||
Db        515 TGAAGAAGGACAACAGAACCTAAAGCCTGCAGCTAACCAATAGAAATTACTTAAAACCCT 456

Qy       4528 TCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTAC 4587
             ||| |||||||||||||||||| ||||||||||||||:|||||||||||||||||||||
Db        455 TCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTAC 396

Qy       4588 TGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARAR 4647
             |||||||||||||||||||||||||||||:|||:|||||||||||||:|||||||:|:
Db        395 TGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAAAG 336

Qy       4648 AAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCC 4707
             ||||||||||||||:||| |||||||||||||:|| |||:|||||||||||||||||
```

```
Db        335 AAATAATCCCCTGCCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCC 276

Qy       4708 TGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCT 4766
              |:|:||:|||||||||| |||||||||||| ||||||| |||| ||||||||||||||
Db        275 AGGAGAAGACTGGCAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCT 217

Qy       4767 ACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAA 4826
              ||||||  || ||:||||||| || ||||||| |:|||||||||| ||| |||| |||||||
Db        216 ACTAGTTCGGCTAGATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAA 157

Qy       4827 GGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCG 4886
              ||||| |||| |||||||| |||||||||||||||||||||||||||| ||||| |||| 
Db        156 GGCCCGAGAGATAATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCG 98

Qy       4887 AGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCA 4946
              |||||||||||||||||| |||| |||||| |||| ||||||||||||||||||||||||
Db         97 AGGCTTACAGAGTGACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCA 38

Qy       4947 GGCGTTAGGTATACGATATCACTTACACTGCGCCTG 4982
              ||| ||||| |||| |||||||||||||| |||||||
Db         37 GGCATTAGGCATACAATATCACTTACACCGCGCCTG 2


RESULT 66
US-10-301-480-534555/c
; Sequence 534555, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 534555
;   LENGTH: 990
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-534555

  Query Match            10.4%;  Score 776.6;  DB 1;  Length 990;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  864;  Conservative   42;  Mismatches   77;  Indels   14;  Gaps    6;

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
              |||||||||||||||| |||:||||:|||||||||||||||:|||:|  ||||||||||
Db        990 CTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCAGGCACTAGAA 931

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
              ||||| || |:||||| ||:|||||||||||||| |||:|:|||||:|||||||| ||
Db        930 TTAGGGGACGGAAAAATGGTAAATATATATACAGAATCTAAGTATGCTTACCTAGTCCTC 871

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
              || |||||||::||||||||:|:|||||||||||||||||||||:|||:||||||||||
Db        870 CACGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGGAACACCTATC 811

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
              |:|||||||||||||||||||:|||||||:||||: || |||||||||:||||||||:|
Db        810 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCTATAGAGAAACCTAAAGAGGTGGCA 751

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
              ||||||||||||:|||||||||     ||||||:||||||||||:|||:|||: |||||
Db        750 GTCTTACACTGCTGGGGTCATC-----GGAAAGGAAAGGGAAATAGAAGGGAACCGCCAA 696

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
              |  :||||||||||:|||||||| |||||||||| |||||||||||||| |||||  ||
Db        695 GTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACACTCCATTAGAAATCCTTATAGAA 636

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
              ||| |||||| |||||||||| || |||||||||||||||||||||| ||||||||
Db        635 GGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCTGCAGGAGAC 576

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
              | |||||||||||||||||| |||     ||||| ||| |||| ||| || ||||||||||
Db        575 ATAGAATGGGGAACCTCACAAGGACATAGTTTCCTCTCCTCAGGATGGCTAGCCACTGAA 516

Qy       4475 GAAGGGAAA---ATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCAT 4531
              |||||  ||    || |  ||||||| ||| ||||| ||||||||||||||||||||
Db        515 GAAGGACAACAGCTAATAAAGCCTGCAGCTAACCAATAGAAATTACTTAAAACCCTTCAC 456

Qy       4532 CAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGA 4591
              |||||||||||||||||| ||||||||||||||:|||||||||||||||||||||||||
Db        455 CAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGA 396

Qy       4592 CCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAAT 4651
              |||||||||||||||||||||||||:|||:|||||||||||||:|||||||:|:||||
Db        395 CCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAAAGAAAT 336

Qy       4652 AATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRA 4711
              |||||||||||||:||| ||||||||||||||:|| |||:|||||||||||||| |:|
Db        335 AATCCCCTGCCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGA 276

Qy       4712 RAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTA 4770
              :||:|||||||||| |||||||||||| ||||||| |||| ||||||||||||||||||
Db        275 GAAGACTGGCAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTA 217

Qy       4771 GTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCC 4830
              ||  || ||:||||||| || ||||||| |:|||||||||| ||| |||| ||||||||||
Db        216 GTTCGGCTAGATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCC 157

Qy       4831 CAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGC 4890
              | |||| ||||||||  ||||||||||||||||||||||||||||||| ||||| ||||||||
Db        156 CGAGAGATAATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGC 98

Qy       4891 TTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCG 4950
              |||||||||||||| |||| |||||| |||| ||||||||||||||||||||||||||
Db         97 TTACAGAGTGACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCA 38

Qy       4951 TTAGGTATACGATATCACTTACACTGCGCCTGAAGGC 4987
              ||||| |||| |||||||||||||| ||||||| ||||
Db         37 TTAGGCATACAATATCACTTACACCGCGCCTGGAGGC 1


RESULT 67
US-10-301-480-1147964/c
; Sequence 1147964, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1147964
;   LENGTH: 990
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1147964

  Query Match            10.4%;  Score 776.6;  DB 1;  Length 990;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  864;  Conservative   42;  Mismatches   77;  Indels   14;  Gaps    6;
```

```
Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
Db        990 CTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCAGGCACTAGAA 931

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
Db        930 TTAGGGGACGGAAAAATGGTAAATATATATACAGAATCTAAGTATGCTTACCTAGTCCTC 871

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
Db        870 CACGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGGAACACCTATC 811

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
Db        810 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCTATAGAGAAACCTAAAGAGGTGGCA 751

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
Db        750 GTCTTACACTGCTGGGGTCATC-----GGAAAGGAAAGGGAAATAGAAGGGAACCGCCAA 696

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
Db        695 GTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACACTCCATTAGAAATCCTTATAGAA 636

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
Db        635 GGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCTGCAGGAGAC 576

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
Db        575 ATAGAATGGGGAACCTCACAAGGACATAGTTTCCTCTCCTCAGGATGGCTAGCCACTGAA 516

Qy       4475 GAAGGGAAA---ATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCAT 4531
Db        515 GAAGGACAACAGCTAATAAAGCCTGCAGCTAACCAATAGAAATTACTTAAAACCCTTCAC 456

Qy       4532 CAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGA 4591
Db        455 CAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGA 396

Qy       4592 CCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAAT 4651
Db        395 CCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAAAGAAAT 336

Qy       4652 AATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRA 4711
Db        335 AATCCCCTGCCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGA 276

Qy       4712 RAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTA 4770
Db        275 GAAGACTGGCAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTA 217

Qy       4771 GTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCC 4830
Db        216 GTTCGGCTAGATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCC 157

Qy       4831 CAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGC 4890
Db        156 CGAGAGATAATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGC 98

Qy       4891 TTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCG 4950
Db          97 TTACAGAGTGACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCA 38

Qy       4951 TTAGGTATACGATATCACTTACACTGCGCCTGAAGGC 4987
Db          37 TTAGGCATACAATATCACTTACACCGCGCCTGGAGGC 1
```

RESULT 68
US-09-925-065A-918077
; Sequence 918077, Application US/09925065A
; Publication No. US20040181048A1

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918077
; LENGTH: 878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-918077

Query Match 10.2%; Score 760.3; DB 1; Length 878;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 762; Conservative 74; Mismatches 41; Indels 3; Gaps 2;

```
Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
Db          2 TGCTTTCTCAAATACCAGAGGAAGCAGAATGGTTTACAGTCCTGGATATTAAGGATGCCT 61

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
Db         62 TTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCGA 121

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
Db        122 ATGCAACATCTCAACTCACCTGGACTTTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 181

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
Db        182 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACACTCTTCTCC 241

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
Db        242 TTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACCTTGTGCCATCAAGCCA 301

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
Db        302 CCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGACCC 361

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
Db        362 AGCTCTGCTCACAGCAGKTTAAATACTTGGGGCTAAAATTATCCAAAGGCACCAGGGCCC 421

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
Db        422 TCAGTGAGGAACATATCAAGCCTATACTGGCTTATCCTCATCCCCAAATCCTAAAGCAAC 481

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
Db        482 TAAGAGAGTTCCTTGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAGGTATGGCAAA 540

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
Db        541 ATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 600

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
Db        601 AGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGAGGGCCCTAACCCAAG 658
```

```
Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
Db        659 CCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTATGTCACAGAAAAAACAG 718

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
Db        719 GAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTTGCAACCCATGGCATAC 778

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTG 3585
Db        779 CTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCG 838

Qy       3586 GCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
Db        839 GCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATAC 878


RESULT 69
US-09-925-065A-918077
; Sequence 918077, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918077
;   LENGTH: 878
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-918077

  Query Match             10.2%;  Score 760.3;  DB 1;  Length 878;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  762;  Conservative   74;  Mismatches   41;  Indels    3;  Gaps    2;

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
Db          2 TGCTTTCTCAAATACCAGAGGAAGCAGAATGGTTTACAGTCCTGGATATTAAGGATGCCT 61

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
Db         62 TTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCGA 121

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
Db        122 ATGCAACATCTCAACTCACCTGGACTTTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 181

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
Db        182 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACACTCTTCTCC 241

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
Db        242 TTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACCTTGTGCCATCAAGCCA 301

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
Db        302 CCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGACCC 361

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
Db        362 AGCTCTGCTCACAGCAGKTTAAATACTTGGGGCTAAAATTATCCAAAGGCACCAGGGCCC 421

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
Db        422 TCAGTGAGGAACATATCAAGCCTATACTGGCTTATCCTCATCCCCAAATCCTAAAGCAAC 481

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
Db        482 TAAGAGAGTTCCTTGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAGGTATGGCAAA 540

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
Db        541 ATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 600

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
Db        601 AGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGAGGGCCCTAACCCAAG 658

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
Db        659 CCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTATGTCACAGAAAAAACAG 718

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
Db        719 GAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTTGCAACCCATGGCATAC 778

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTG 3585
Db        779 CTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCG 838

Qy       3586 GCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
Db        839 GCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATAC 878


RESULT 70
US-10-301-480-567825/c
; Sequence 567825, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 567825
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-567825

  Query Match             9.8%;  Score 733.101;  DB 1;  Length 1000;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  828;  Conservative   46;  Mismatches   93;  Indels   17;  Gaps    6;

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
Db        997 AACATTTGCTTAAATATCAGGCCCTATTACTTGAAGGACCAGTGCTGCGACTGCTCACTT 938

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
Db        937 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGTAAAGATGGAACATAACC 878
```

```
Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||    |||| ||||||||||   |  |||||||||||||:|| |:||| |: |||
Db        877 GTCAAC----AATTGCTCAAACCTACACTGCTCGAGGGGACCTTCTAAAGGTTGCCCTGA 822

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| |||||||||||||||||||||| |||||||||||||||||||||| ||
Db        821 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGCTCCTTTGTAGAAAAAGGACTTTGA 763

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGG----AATAYTTGAAAGTAATCCCCTCAC 4002
             || |: |||||||||||||||||||||||||    ||||:|||||||||| ||||||||
Db        762 AAGGCAGGGTATGCAGTGGTCAGTGATAATGGAATAAATACTTGAAAGTAAGCCCCTCAC 703

Qy       4003 TCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGG 4062
             |||||||||||| |||:||||:|||||||||||||||||:|||:| ||||||||||||||
Db        702 TCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAATTAGG 643

Qy       4063 AGAAGRAAAAAGGGYAA--ATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
             |||||:||| |||   ||  |||||||||||||||||:|: || ||:|||||||| |||||
Db        642 AGAAGGAAAGAGGAAAAATATATATATACAGACTCTAAGAATTCTTACCTAGTCCTCCAT 583

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMA 4180
             |||||||::|||||||||:|:|||||||||||||||||||||:|||:|||| |||||:|
Db        582 GCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAACATCTATCAAA 523

Qy       4181 CATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTC 4240
             |||||||||||| |||||:|||||   ||||:|||||||||||||||:|||||||:||||
Db        522 CATCAGGAAGCCACTAGGAGATTAT---TGGCTGTACAGAAACCTAAAGAGGTGGCAGTC 466

Qy       4241 TTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCA 4300
             |||||||||:||||| |||| |||||||||:||| ||||| |:|||: ||:  || ||||
Db        465 TTACACTGCTGGGGTTATCAGAAAGGAAAGGAAAAGGAAACAGAAGAAAACCACCGAGCA 406

Qy       4301 KATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTT 4360
             :|||||||| |:||||||||  |||||  |||||||| ||||||||||||||  ||
Db        405 GATATTGAACCCAAAAGAGCCACAAGGTGGGACCCTCCTTTAGAAATGCTTATAGAAGGA 346

Qy       4361 CCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACA 4420
             ||| |||||| |||||||||| || ||||||||||||||||||||||||||||| ||||| |
Db        345 CCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATA 286

Qy       4421 GAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAA 4477
             |||| || |||||||||||     || || |||||||| ||| || ||||||||||||||
Db        285 GAATAGGTAACCTCACGAGGACATAGCTTCCTCCCCTCAGGATGGCTAGCCACTGAAGAA 226

Qy       4478 GGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACC 4537
             || |||||||||||| ||||| ||| |||||||||||||||||||| ||| | ||||||
Db        225 GGAAAAATACTTTTGTCTGCAGCTAAGCAATGGAAATTACTTAAAACTCTTAACCAAACC 166

Qy       4538 TTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGC 4597
             |||||||||||||| |||||||||||| |:|||||||||||| ||||||||||||||||
Db        165 TTTCACTTAGGCATTGATAGCACCCATGAGATGGCCAAATCATAATTTACTGGACCAGGC 106

Qy       4598 CTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCC 4657
              |||||||||||||||||:|||:|||||||||| ||:|||||||:|:|||||||||
Db        105 ATTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTAAAGTGTGCCAAAGAAATAATCCT 46

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCA 4681
             ||||  :   | | |     ||||
Db         45 CTGCACTTCAGGCCATACATTTCA 22
```

```
RESULT 71
US-10-301-480-1181234/c
; Sequence 1181234, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1181234
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1181234

  Query Match             9.8%;  Score 733.101;  DB 1;  Length 1000;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  828;  Conservative   46;  Mismatches   93;  Indels   17;  Gaps    6;

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
             | : ||| |||||||:|||||| |||||||||||:| |||||||||:|||| |||||
Db        997 AACATTTGCTTAAATATCAGGCCCTATTACTTGAAGGACCAGTGCTGCGACTGCTCACTT 938

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
             |||||||||||||||||:| |||||||||||||||||||| |||||| :||:|||||
Db        937 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGTAAAGATGGAACATAACC 878

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||    |||| ||||||||||   |  |||||||||||||:|| |:||| |: |||
Db        877 GTCAAC----AATTGCTCAAACCTACACTGCTCGAGGGGACCTTCTAAAGGTTGCCCTGA 822

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| |||||||||||||||||||||| |||||||||||||||||||||| ||
Db        821 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGCTCCTTTGTAGAAAAAGGACTTTGA 763

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGG----AATAYTTGAAAGTAATCCCCTCAC 4002
             || |: |||||||||||||||||||||||||    ||||:|||||||||| ||||||||
Db        762 AAGGCAGGGTATGCAGTGGTCAGTGATAATGGAATAAATACTTGAAAGTAAGCCCCTCAC 703

Qy       4003 TCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGG 4062
             |||||||||||| |||:||||:|||||||||||||||||:|||:| ||||||||||||||
Db        702 TCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAATTAGG 643

Qy       4063 AGAAGRAAAAAGGGYAA--ATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
             |||||:||| |||   ||  |||||||||||||||||:|: || ||:|||||||| |||||
Db        642 AGAAGGAAAGAGGAAAAATATATATATACAGACTCTAAGAATTCTTACCTAGTCCTCCAT 583

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMA 4180
             |||||||::|||||||||:|:|||||||||||||||||||||:|||:|||| |||||:|
Db        582 GCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAACATCTATCAAA 523

Qy       4181 CATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTC 4240
             |||||||||||| |||||:|||||   ||||:|||||||||||||||:|||||||:||||
Db        522 CATCAGGAAGCCACTAGGAGATTAT---TGGCTGTACAGAAACCTAAAGAGGTGGCAGTC 466

Qy       4241 TTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCA 4300
             |||||||||:||||| |||| |||||||||:||| ||||| |:|||: ||:  || ||||
Db        465 TTACACTGCTGGGGTTATCAGAAAGGAAAGGAAAAGGAAACAGAAGAAAACCACCGAGCA 406

Qy       4301 KATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTT 4360
             :|||||||| |:||||||||  |||||  |||||||| ||||||||||||||  ||
Db        405 GATATTGAACCCAAAAGAGCCACAAGGTGGGACCCTCCTTTAGAAATGCTTATAGAAGGA 346

Qy       4361 CCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACA 4420
             ||| |||||| |||||||||| || ||||||||||||||||||||||||||||| ||||| |
Db        345 CCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATA 286

Qy       4421 GAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAA 4477
             |||| || |||||||||||     || || |||||||| ||| || ||||||||||||||
Db        285 GAATAGGTAACCTCACGAGGACATAGCTTCCTCCCCTCAGGATGGCTAGCCACTGAAGAA 226

Qy       4478 GGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACC 4537
             || |||||||||||| ||||| ||| |||||||||||||||||||| ||| | ||||||
Db        225 GGAAAAATACTTTTGTCTGCAGCTAAGCAATGGAAATTACTTAAAACTCTTAACCAAACC 166

Qy       4538 TTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGC 4597
             |||||||||||||| |||||||||||| |:|||||||||||| ||||||||||||||||
```

```
Db        165 TTTCACTTAGGCATTGATAGCACCCATGAGATGGCCAAATCATAATTTACTGGACCAGGC 106

Qy       4598 CTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCC 4657
             ||||||||||||||||||:|||:|||||||||| ||:|||||||:|:|||||||||
Db        105 ATTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTAAAGTGTGCCAAAGAAATAATCCT 46

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCA 4681
             ||||  :   | | |     ||||
Db         45 CTGCACTTCAGGCCATACATTTCA 22


RESULT 72
US-10-450-763-7529
; Sequence 7529, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 7529
;   LENGTH: 910
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (28)..(909)
;   OTHER INFORMATION: 73% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1108.
US-10-450-763-7529

  Query Match             9.5%;  Score 714.8;  DB 1;  Length 910;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  821;  Conservative   25;  Mismatches   52;  Indels   40;  Gaps    7;

Qy       2238 TTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 2297
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 TTCTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAAT 65

Qy       2298 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 2357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCAT 125

Qy       2358 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAA 2417
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        126 TATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAA 184

Qy       2418 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCT 2477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCT 244

Qy       2478 GTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGA 2537
             |||||||||||||||||||  | ||||||||||||||||||| ||||||||  |||
Db        245 GTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGCAATACCCCCTTAGA 304

Qy       2538 CCCAAGGCCCAACAA--GGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
             ||  ||||||||||  || |||||||||||||||||||| |||||||||||||| |||
Db        305 CCTGAGGCCCAACAAAGGGCTTCCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 364

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
             ||||||||||||  |  |||||||||| | |        |||   |  | ||||||||||
Db        365 TAAAACCATGCAATAGCCCCTGCAGTACTCC-------AATTTTAGGAGTACAGAAACCC 417

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
             |  |||||||||| |||||||||||||||||||||||| || |||||| |||||| |
Db        418 AACGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCGAT-GAGGCCATTGTCCCTC 475

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
             ||||||||||||||||||  |||||||||| ||::|||||||||||||||||||||||||
Db        476 TATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGT 535

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
             |||||||:|||||||||||:|||||||||| |||||||||||||||||||||| ||||
Db        536 TGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGAGTCTC 595

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
             |||||||||||||||||||||||| |||||| ||||:  |||||||||||||||||:|||
Db        596 AATTCTTGTTTGCCTTTGAAGATCCTTCAATCCCAATGTCTCAACTCACCTGGACTGTTT 655

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
             |||||||||||||||||||||:|||||||||||  |||  ||||||||||||||||||
Db        656 TACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTCACCATACATTAGCCCAAGACTTGA 715

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
             |:||:|::|||||||||||||||||||||||||                || ||:
Db        716 GCCAGTTCTCATACCTGGACACTCTTGTCC----------------------TTTGCCT 752

Qy       3016 CCYRTTCAGAAACC-----TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGC 3070
             ||::|| |||||||       |||||||||||||||||||| :|||||:||| |||||||
Db        753 CCCGTTGAGAAACCAACAGGAGTGCCATCAAGCCACCCAAGTGCTCTTAAATGTCCTCGC 812

Qy       3071 YACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATA 3130
             :|||||||||||||:|||||||||||| |:|:|||||:|||||||||||||||||||||
Db        813 TACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATA 872

Qy       3131 CTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
             ||||||: |||:|||||||||||| ||||:||||||||
Db        873 CTTAGGGGTAAAATTATCCAAAGGGACCAGGGCCCTCA 910


RESULT 73
US-09-864-761-21192
; Sequence 21192, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21192
;   LENGTH: 792
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC002346.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  20
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
;   OTHER INFORMATION: NT HIT: AF208161.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 2.00e-16
US-09-864-761-21192

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||||||| |||||||||||  ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 74
US-09-864-761-14911
; Sequence 14911, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
```

```
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14911
;   LENGTH: 944
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007567.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
US-09-864-761-14911

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
Db        941 CGAA 944


RESULT 75
US-10-450-763-8028
; Sequence 8028, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 8028
;   LENGTH: 924
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (19)..(924)
;   OTHER INFORMATION: 77% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1201.
US-10-450-763-8028

  Query Match             8.8%;  Score 662.8;  DB 1;  Length 924;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  724;  Conservative   44;  Mismatches   71;  Indels   10;  Gaps    4;

Qy       2494 AGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAG 2553
Db         86 AGGACCACTTTCTTTGTGGTCAAGAAAGGCAGTACCCCCTTAGACCCAAGGCCCAACAAG 145

Qy       2554 GATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTC 2613
Db        146 GACTCCAAAAAACTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCC 205

Qy       2614 CCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTT 2673
Db        206 CCTGCAATACTCC-------AACTGTAGGAGTACAGAAACCCAACGGACAGTGGA-GGTT 257

Qy       2674 AGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTA 2733
Db        258 AGTGCAAAATCTCAGGATTATCAAT-GAGACTGTTGTTCCTCTATACCCAGCTGAACCTA 316

Qy       2734 GCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACC 2793
Db        317 ACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTATATAGTCCTGGACC 376

Qy       2794 TTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTG 2853
Db        377 TTAAGGATGCCTTCTTCTGCATCCCTGTACATCCTGATTCCCAATTCCTGTTTGCCTTTG 436

Qy       2854 AAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
```

```
                |||||  ||||||||  ||:|  ||||||||||||||||||:||||||  ||||||||||||||
Db        437 AAGATCCTTCAAACTCAACGTCTCAACTCACCTGGACTGTTTTACGCCAAGGGTTCAGGG 496

Qy       2914 ATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGG 2973
                ||||:||||||||||||||||||||||||||||||  |||||:||:|::|||||||||||
Db        497 ATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGATTTGAGCCAATTCTCATACCTGG 556

Qy       2974 ACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGT 3033
                |||||||||||||| :|||:|||||||||||||||||:||: ||::|||||||||||||||
Db        557 ACACTCTTGTCCTTTGGTATGTGGATGATTTACTTTTAGCCACCTGTTCAGAAACCTTGT 616

Qy       3034 GCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCA 3093
                |||||||||||||||||  :|||||:|||||||||  |:|||||||||||||:||||||||
Db        617 GCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCA 676

Qy       3094 AACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAG 3153
                |||:|:|:|||||:|||||||||||||||||||||||||||||||:||||:||||||||||
Db        677 AACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAG 736

Qy       3154 GCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAA 3213
                ||    |:||||||||||||||||:::|||||||||||||||||||||||||  ||:|||||
Db        737 GCTTGAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCTTCCCAAAA 796

Qy       3214 CCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCC 3273
                |||||||||||||||:|::|||||||||:|||:|||:|||| |||||: |||||| |||
Db        797 CCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTACCGAATGTGGATT-CCC 855

Qy       3274 AGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAAT 3333
                ||||:  ||:|||||| ||| :|||||:|||||:||||||| ||||||||||||||||||
Db        856 AGGTACAGCAAAATAGGCAGATCATTATATACACTAATTAAAGAAACTCAGAAAGCCAAT 915

Qy       3334 ACCCATTTA 3342
                |||||||||
Db        916 ACCCATTTA 924


RESULT 76
US-09-925-065A-923664/c
; Sequence 923664, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923664
;   LENGTH: 799
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-923664

  Query Match             8.7%;  Score 651.8;  DB 1;  Length 799;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  700;  Conservative   58;  Mismatches   41;  Indels   18;  Gaps    5;

Qy       3186 CCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTA 3245
                |||| |||||||||||||||:|||||||||||||||||||:| :|||||||||:||
Db        799 CCTACACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGCGTTCCTTGGCATA 740

Qy       3246 AYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATAC 3305
                |:|||:|||||| |||:|||||||:||| :|||||||||| :|||||:||||
Db        739 ACAGGTTTCTGCGGAATATGGATT-CCCAGGTATGGTGAAATAGCCAGACCATTATATAC 681

Qy       3306 ASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRA 3365
                |:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||| |:::|
Db        680 ACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAGGCAGA 621

Qy       3366 AGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCA 3425
                ||||||||||||||||  ||  | | |||  ||||||||:|  ||||||||||:||||||
Db        620 AGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCACAGTGTTAAGCTTGCCA 563

Qy       3426 ACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCT 3485
                | :|| ||||||||||:||:||| :|||||||||||  ||:||:||||||:| ||||||
Db        562 AGGGGACAAGACTTTTCTTTATAAGTCACAGAAAAAA--AGGAATAGCTCTAGTAGTCCT 505

Qy       3486 TACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGT 3545
                ||| |||:|||:|||||  ||||||||||||::||||::||||||:||||||||:||||||
Db        504 TACTCAGGTCCGAGGGAAGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAATTGATGT 445

Qy       3546 AGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATC 3605
                |||||||||||||||:|:||| ||||||: ||||||| || |||||||| ::|||||||
Db        444 AGTGGCAAAGGGTTGGCCTCACTGTTTACAGGTAGTGCTGACAGTAGCAGGCTTAGTATC 385

Qy       3606 TGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAA 3665
                ||||||  ||||||||||||||||||:|||||||||||||||||||||||||||||:|||||
Db        384 TGAAGCCGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAA 325

Qy       3666 YRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCA 3725
                ::|||||||||||||||||||||||||||||||||||||||||||||            ||||:|||
Db        324 CAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGAC-----------AAATATCA 277

Qy       3726 GGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGY 3785
                ||||||||||||||:||||||||||||:||||  ||| ||||||||||| |||||||:
Db        276 GGCTCTATTACTTGAAAGGCCAGTGCTGCGACTGCACACATGTGCAACTCTCAACCCAGC 217

Qy       3786 CNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCA 3845
                |  |||||||||||||||||||||||||||| :||:|||| |||||||||: |||| ||||
Db        216 CAAATTTCTTCCAGACAATGAAGAAAAGATGGAACATAATTGTCAACAAGCAATTGCTCA 157

Qy       3846 AACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACT 3905
                ||||||:|||||| || ||||||||:||||:|||||:||||||||||:|||| ||||||
Db        156 AACCTAYGCCACTTGAAGGGACCTTTTAGAGGTTCCCTTGACTGATCCTGACC-TCAACT 98

Qy       3906 TGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGG 3965
                |||||||||||||||||||||||| |||||||||||||| |||||:||||||||||| |
Db         97 TGTATACTGATGGAAGTTCCTTTGCAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTTG 38

Qy       3966 TCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCAC 4002
                |||||||||||||||||:|||||||||||||||||| |
Db         37 TCAGTGATAATGGAATACTTGAAAGTAATCCCCTCCC 1


RESULT 77
US-09-925-065A-923664/c
; Sequence 923664, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923664
;    LENGTH: 799
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-923664

  Query Match             8.7%;  Score 651.8;  DB 1;  Length 799;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  700;  Conservative   58;  Mismatches   41;  Indels   18;  Gaps    5;

Qy       3186 CCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTA 3245
             |||| |||||||||||||||||:|||||||||||||||||||||:| :|||||||||:||
Db        799 CCTACACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGCGTTCCTTGGCATA 740

Qy       3246 AYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATAC 3305
             |:|||:|||||| |||:||||||| |||||||:||| :|||||||||| :|||||:||||
Db        739 ACAGGTTTCTGCGGAATATGGATT-CCCAGGTATGGTGAAATAGCCAGACCATTATATAC 681

Qy       3306 ASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRA 3365
             |:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||| |:::|
Db        680 ACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAGGCAGA 621

Qy       3366 AGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCA 3425
             |||||||||||||||||  ||  | | |||. ||||||||:| ||||||||||:||||||
Db        620 AGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCACAGTGTTAAGCTTGCCA 563

Qy       3426 ACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCT 3485
             | :|| |||||||||||:||:||| :||||||||||||  ||:||:||||||:| ||||||
Db        562 AGGGGACAAGACTTTTCTTTATAAGTCACAGAAAAAA--AGGAATAGCTCTAGTAGTCCT 505

Qy       3486 TACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGT 3545
             ||| |||:|||:|||||| |||||||||||::||||::|||||||:|||||||:|||||
Db        504 TACTCAGGTCCGAGGGAAGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAATTGATGT 445

Qy       3546 AGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATC 3605
             |||||||||||||||:|:||| |||||||: ||||||| || |||||||||| ::|||||||
Db        444 AGTGGCAAAGGGTTGGCCTCACTGTTTACAGGTAGTGCTGACAGTAGCAGGCTTAGTATC 385

Qy       3606 TGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAA 3665
             |||||| |||||||||||||||||:|||||||||||||||||||||||||||||:|||||
Db        384 TGAAGCCGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAA 325

Qy       3666 YRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCA 3725
             ::|||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        324 CAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGAC-----------AAATATCA 277

Qy       3726 GGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGY 3785
             |||||||||||||||:|||||||||||||:|||| ||| |||||||||| |||||||:
Db        276 GGCTCTATTACTTGAAAGGCCAGTGCTGCGACTGCACACATGTGCAACTCTCAACCCAGC 217

Qy       3786 CNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCA 3845
             | |||||||||||||||||||||||||||| :||:|||| |||||||||: |||| ||||
Db        216 CAAATTTCTTCCAGACAATGAAGAAAAGATGGAACATAATTGTCAACAAGCAATTGCTCA 157

Qy       3846 AACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACT 3905
             ||||||:|||||| || |||||||||:||||:|||||:||||||||||:|||| ||||||
Db        156 AACCTAYGCCACTTGAAGGGACCTTTTAGAGGTTCCCTTGACTGATCCTGACC-TCAACT 98

Qy       3906 TGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGG 3965
             |||||||||||||||||||||||| |||||||||||||| |||||:||||||||||| |
Db          97 TGTATACTGATGGAAGTTCCTTTGCAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTTG 38

Qy       3966 TCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCAC 4002
             ||||||||||||||||:|||||||||||||||| |
Db          37 TCAGTGATAATGGAATACTTGAAAGTAATCCCCTCCC 1


RESULT 78
US-09-925-065A-941867
; Sequence 941867, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941867
;    LENGTH: 795
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-941867

  Query Match             8.7%;  Score 649.4;  DB 1;  Length 795;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  697;  Conservative   58;  Mismatches   40;  Indels   18;  Gaps    5;

Qy       3181 TCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTG 3240
             ||||||||| |||||||||||||||||:|||||||||||||||||||||||:| :|||||||
Db          1 TCCAGCCTACACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGCGTTCCTTG 60

Qy       3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300
             ||:|||:|||:|||||| |||:||||||| |||||||:||| :|||||||||| :|||||
Db         61 GCATAACAGGTTTCTGCGGAATATGGATT-CCCAGGTATGGTGAAATAGCCAGACCATTA 119

Qy       3301 WATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAA 3360
             :|||||:|||||||||||||||||||||||||||||||||||||:|||||||||:|:||||
Db        120 TATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAG 179

Qy       3361 GYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420
             |:::|||||||||||||||  || | | |||  ||||||||:| ||||||||||:|
Db        180 GCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCACAGTGTTAAGCT 237

Qy       3421 TGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGA 3480
             |||||| :|| ||||||||||:||:||| :|||||||||||  ||:||:||||||:| |
Db        238 TGCCAAGGGGACAAGACTTTTCTTTATAAGTCACAGAAAAAA--AGGAATAGCTCTAGTA 295

Qy       3481 GTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540
             |||||||| |||:|||:||||| ||||||||||||::||||::|||||:||||||||:|
Db        296 GTCCTTACTCAGGTCCGAGGGAAGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAATT 355

Qy       3541 GATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600
             |||||||||||||||||||:|:||| |||||||: ||||||| || |||||||||| ::||
Db        356 GATGTAGTGGCAAAGGGTTGGCCTCACTGTTTACAGGTAGTGCTGACAGTAGCAGGCTTA 415

Qy       3601 GTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660
             |||||||||| ||||||||||||||||:|||||||||||||||||||||||||||||:
Db        416 GTATCTGAAGCCGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGAT 475

Qy       3661 GTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAA 3720
             |||||::|||||||||||||||||||||||||||||||||||||||||||||  |||
Db        476 GTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGAC-----------AAA 523

Qy       3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAAC 3780
             |:|||||||||||||||||:|||||||||||||:|||| ||| |||||||||| |||
Db        524 TATCAGGCTCTATTACTTGAAAGGCCAGTGCTGCGACTGCACACATGTGCAACTCTCAAC 583

Qy       3781 CCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATT 3840
             ||||:|  |||||||||||||||||||||||| :||:|||| |||||||||: ||||
Db        584 CCAGCCAAATTTCTTCCAGACAATGAAGAAAAGATGGAACATAATTGTCAACAAGCAATT 643
```

```
Qy       3841 TCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTT 3900
Db        644 GCTCAAACCTAYGCCACTTGAAGGGACCTTTTAGAGGTTCCCTTGACTGATCCTGACC-T 702

Qy       3901 CAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960
Db        703 CAACTTGTATACTGATGGAAGTTCCTTTGCAGAAAAAGGACTTCAAAAAGTGGGGTATGC 762

Qy       3961 AGTGGTCAGTGATAATGGAATAYTTGAAAGTAA 3993
Db        763 AGTTGTCAGTGATAATGGAATACTTGAAAGTAA 795


RESULT 79
US-09-925-065A-941867
; Sequence 941867, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941867
;   LENGTH: 795
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-941867

  Query Match             8.7%;  Score 649.4;  DB 1;  Length 795;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  697;  Conservative   58;  Mismatches   40;  Indels   18;  Gaps    5;

Qy       3181 TCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTG 3240
Db          1 TCCAGCCTACACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGCGTTCCTTG 60

Qy       3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300
Db         61 GCATAACAGGTTTCTGCGGAATATGGATT-CCCAGGTATGGTGAAATAGCCAGACCATTA 119

Qy       3301 WATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAA 3360
Db        120 TATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAG 179

Qy       3361 GYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYT 3420
Db        180 GCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCACAGTGTTAAGCT 237

Qy       3421 TGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGA 3480
Db        238 TGCCAAGGGGACAAGACTTTTCTTTATAAGTCACAGAAAAAA--AGGAATAGCTCTAGTA 295

Qy       3481 GTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540
Db        296 GTCCTTACTCAGGTCCGAGGGAAGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAATT 355

Qy       3541 GATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600
Db        356 GATGTAGTGGCAAAGGGTTGGCCTCACTGTTTACAGGTAGTGCTGACAGTAGCAGGCTTA 415

Qy       3601 GTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660
Db        416 GTATCTGAAGCCGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGAT 475

Qy       3661 GTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAA 3720
Db        476 GTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGAC------------AAA 523

Qy       3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAAC 3780
Db        524 TATCAGGCTCTATTACTTGAAAGGCCAGTGCTGCGACTGCACACATGTGCAACTCTCAAC 583

Qy       3781 CCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATT 3840
Db        584 CCAGCCAAATTTCTTCCAGACAATGAAGAAAAGATGGAACATAATTGTCAACAAGCAATT 643

Qy       3841 TCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTT 3900
Db        644 GCTCAAACCTAYGCCACTTGAAGGGACCTTTTAGAGGTTCCCTTGACTGATCCTGACC-T 702

Qy       3901 CAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960
Db        703 CAACTTGTATACTGATGGAAGTTCCTTTGCAGAAAAAGGACTTCAAAAAGTGGGGTATGC 762

Qy       3961 AGTGGTCAGTGATAATGGAATAYTTGAAAGTAA 3993
Db        763 AGTTGTCAGTGATAATGGAATACTTGAAAGTAA 795


RESULT 80
US-09-925-065A-956235/c
; Sequence 956235, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956235
;   LENGTH: 758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956235

  Query Match             8.6%;  Score 646.8;  DB 1;  Length 758;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  662;  Conservative   62;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2739 TATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAG 2798
Db        758 TACACTCTGCTTTCTCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGGCCTTAAG 699

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
Db        698 GATGMCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAA 639

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
```

```
Db        638 CCTTCAAACCCCACATCTCAACTCACCTGGACTGTTTTACCGAAAGGGTTCAGGGATAGC 579

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACT 2978
Db        578 CCCCATCTATTTGGCCAGGTATTAGCCCAAAACTTGAGCCAATTCTCATACCTGGACACT 519

Qy       2979 CTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCAT 3038
Db        518 CTTGTCCTTCGGTACGTGGATGATTTACTTTTAGCTGCCCGTTCAGAAACCTTGTGCCAT 459

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
Db        458 CAAGCCACCCAAGCACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 399

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
Db        398 AAGGCTCAGCTCTGCTCACAGGAGGTTACGTACTTAGGGCTAAAATTATCCAAAGGCACC 339

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
Db        338 AGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTAATCTCATCCTAAAACCCTA 279

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTW 3278
Db        278 AAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-CCCAGGTA 220

Qy       3279 TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCA 3338
Db        219 AGACGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAAAAAGCCAATACCCA 160

Qy       3339 TTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAA 3398
Db        159 TTTAGTAAGATGGACCCCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGAGGGCCCTA 102

Qy       3399 ACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAA 3458
Db        101 ACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-A 43

Qy       3459 AAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
Db          42 AAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAGG 1
```

RESULT 81
US-09-925-065A-956235/c
; Sequence 956235, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956235
; LENGTH: 758
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-956235

Query Match 8.6%; Score 646.8; DB 1; Length 758;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 662; Conservative 62; Mismatches 34; Indels 4; Gaps 3;

```
Qy       2739 TATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAG 2798
Db        758 TACACTCTGCTTTCTCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGGCCTTAAG 699

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
Db        698 GATGMCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAA 639

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
Db        638 CCTTCAAACCCCACATCTCAACTCACCTGGACTGTTTTACCGAAAGGGTTCAGGGATAGC 579

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACT 2978
Db        578 CCCCATCTATTTGGCCAGGTATTAGCCCAAAACTTGAGCCAATTCTCATACCTGGACACT 519

Qy       2979 CTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCAT 3038
Db        518 CTTGTCCTTCGGTACGTGGATGATTTACTTTTAGCTGCCCGTTCAGAAACCTTGTGCCAT 459

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
Db        458 CAAGCCACCCAAGCACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 399

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
Db        398 AAGGCTCAGCTCTGCTCACAGGAGGTTACGTACTTAGGGCTAAAATTATCCAAAGGCACC 339

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
Db        338 AGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTAATCTCATCCTAAAACCCTA 279

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTW 3278
Db        278 AAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-CCCAGGTA 220

Qy       3279 TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCA 3338
Db        219 AGACGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAAAAAGCCAATACCCA 160

Qy       3339 TTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAA 3398
Db        159 TTTAGTAAGATGGACCCCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGAGGGCCCTA 102

Qy       3399 ACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAA 3458
Db        101 ACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-A 43

Qy       3459 AAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
Db          42 AAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAGG 1
```

RESULT 82
US-10-450-763-7526/c
; Sequence 7526, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23

```
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 7526
;   LENGTH: 902
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (705)..(4)
;   OTHER INFORMATION: 92% homologous to Homo sapiens syncytin precursor,accession
;   OTHER INFORMATION: number AF208161,Smith-Waterman Score=1128.
US-10-450-763-7526

  Query Match             8.6%;  Score 646;  DB 1;  Length 902;
  Best Local Similarity   93.6%;  Pred. No. 0;
  Matches  661;  Conservative   10;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       6165 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 6224
             |||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 GCCACATATTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACAC 647

Qy       6225 CACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCT 6284
             ||||||| |||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        646 CACTTCCATTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCT 587

Qy       6285 CACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 6344
             |||| |||||||||||||||||||| | |||||||||||||||| |||||||| |||||
Db        586 CACCCGTGTAAAATTTAGCAATACTATAGACACAACCAACTCCCAGTGCATCAGATGGGT 527

Qy       6345 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 6404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 AACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTC 467

Qy       6405 AGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCC 6464
             ||||||||  |||||||||||||| ||| ||||||||||||||||||||||||||||||
Db        466 AGCCTATCACTGTTTGAATGGCTTTTCGGAATCTATGTGCTTCCTCTCATTCTTAGTGCA 407

Qy       6465 CCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
             |||:|||:|||||||||||||||||||||||||||||| ||||||| || ||||||||||||||
Db        406 CCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTACCTAAGCCCCGCAA 347

Qy       6525 CAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTAC 6584
             |||||||||||||||||||||||||||||||  |||||||| |||||||    | |||||||
Db        346 CAAAAGAGTACCCATTCTTCCTTTTGTTATCCGAGCAGGAATGCTAGGCAGATTAGGTAC 287

Qy       6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
             ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGCATGGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 227

Qy       6645 TGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 6704
             ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        226 TGGTGACATGGAACGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCT 167

Qy       6705 AGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGG 6764
             |||||||||||| ||||:|||||||||||||||||:|||||| || ||:||||||||
Db        166 AGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCAGAAAGAGGGGG 107

Qy       6765 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCAC 6824
             ||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db        106 AACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCAC 47

Qy       6825 TGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGAR 6870
             ||||||||||:|||||||||:|||| |||||||||| |:|||| |:
Db          46 TGAGAAAGTTAAAGAAATTCGAGATGGAATACAACGCAGAGCATAG 1


RESULT 83
US-10-450-763-12271
; Sequence 12271, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 12271
;   LENGTH: 841
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(840)
;   OTHER INFORMATION: 81% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1214.
US-10-450-763-12271

  Query Match             8.5%;  Score 637.7;  DB 1;  Length 841;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  750;  Conservative   24;  Mismatches   62;  Indels   39;  Gaps    7;

Qy       2299 ATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATT 2358
             |||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATT 60

Qy       2359 ATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAAT 2418
             |||||||||||||||||||||||||||||| ||||||||| ||||||||| ||||||||||
Db          61 ATACACCTGAACATAGGAGAAGGAACACCTGTTTGTTGT-CCCCTGCTTAAGGAAGGAAT 119

Qy       2419 TAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTG 2478
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||| |||||
Db         120 TAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCCATCCTG 178

Qy       2479 TTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGAC 2538
             |||||||||||||||||||| | |||||||||||||||||| |||||||| ||||
Db         179 TTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGCAATACCCCCTTAGAC 238

Qy       2539 CCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAA 2598
             |  |||||||||||||| |||||||||||||||||||| ||||||  |||||| ||||||
Db         239 CTGAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTAAAAGTGCAAGGCCTAGTAA 298

Qy       2599 AACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGT 2658
             || ||||||  |  ||||||  || | |        |||   |  | ||||||||||| |
Db         299 AAACATGCAATAGCCCCTGCGATACTCC-------AATTTTAGGAGTACAGAAACCCAAT 351

Qy       2659 GGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTAT 2718
             ||||||||| |||| |||||||||||||||||| || |||||| |||||| | |||
Db         352 GGACAGTGGA-GGTTGGTGCAAGATCTCAGGATTATCGAT-GAGGCCATTGTCCCTCTAT 409

Qy       2719 ACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGT 2778
             |||||||||||||| |||||||||| ||::|||||||||||||||||||||||||||| ||
Db         410 ACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTTGT 469

Qy       2779 TTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAAT 2838
             |||||:|||||||||||:|||||||||| ||||||||||||||||||||||| |||||||
Db         470 TTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGAGTCTCAAT 529

Qy       2839 TCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTAC 2898
             |||||||||||||||||||| |||||| ||||:  |||||||||||||||||:|||||||
Db         530 TCTTGTTTGCCTTTGAAGATCCTTCAATCCCAATGTCTCAACTCACCTGGACTGTTTTAC 589

Qy       2899 CCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYC 2958
             ||||||||||||||||||:|||||||||||  |||  |||||||||||||||||||:|
Db         590 CCCAAGGGTTCAGGGATAGTCCCCATCTATTTCACCATACATTAGCCCAAGACTTGAGCC 649

Qy       2959 ARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCY 3018
             |:|::|||||||||||||||||||||||                    || ||: ||:
Db         650 AGTTCTCATACCTGGACACTCTTGTCC----------------------TTTGCCTCCC 686
```

```
Qy       3019 RTTCAGAAACC-----TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
Db        687 GTTGAGAAACCAACAGGAGTGCCATCAAGCCACCCAAGTGCTCTTAAATGTCCTCGCTAC 746

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTT 3133
Db        747 CTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTT 806

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCA 3168
Db        807 AGGGGTAAAATTATCCAAAGGGACCAGGGCCCTCA 841


RESULT 84
US-09-925-065A-928581
; Sequence 928581, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928581
;   LENGTH: 825
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928581

  Query Match             8.5%;  Score 637.2;  DB 1;  Length 825;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  677;  Conservative   67;  Mismatches   75;  Indels   10;  Gaps    3;

Qy       2732 TAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGA 2791
Db          1 TAACCCTTATACTCTGCTTTCCCAAATGCCAGAGGAGGCAGAGTGGTTTATAGTCCTGGA 60

Qy       2792 CCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 2851
Db         61 CCTTAAGGATGTCTTTTTCTGCATTCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 120

Qy       2852 TGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAG 2911
Db         121 TGGAGATCCTTCTAACCCAATGTCTCAACTCACCTGGAATGTTTTACCCCAAGGGTTCAG 180

Qy       2912 GGATAGYCCCCATCTATTTGGCCAGGCAT------TAGCCCAAGACTTGAGYCARTYMTC 2965
Db         181 GGATAGCCCTCATCTATTCGGCCAGGCATCAACCCAAGCCCAACACTTGAGCCAGTTCTC 240

Qy       2966 ATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGA 3025
Db         241 ATACTTGAACACTCATGTTGTTTGGTACATGGATGATTTACATCTAGCCGCCCCTGCAGA 300

Qy       3026 AACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
Db         301 AGCCTTGTGCCATCAAGCCACCCAAGTGCTTTTAAACTTCCTCACCACTTGTGGCTACAA 360

Qy       3086 GGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATT 3145
Db         361 GGTTTCCAAACCAAAGTCTCAGCTCTRCTTATAGCAGGTTAAATACTTAGGGTTAAAATT 420

Qy       3146 ATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCA 3205
Db         421 ATCCAAAGGCACCAGGGCCCTCAGCGAGGAATGTATCCATCCTATACTGGCTTATCCTGA 480

Qy       3206 TCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATG 3265
Db         481 TCCCAAAACCCTGAAGCAATTAAGAGGTTTCCATGGCATAACAAGCTTCTGCCAAATACG 540

Qy       3266 GATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGA 3325
Db         541 GATT-CCCAGGTACAGCGAAAGAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGA 599

Qy       3326 AAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
Db         600 AAGCCAATACTCATTTAGTAGAATGGACACCTGAAGCAGAAACAGCTTTTCAGGCCCTAA 659

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
Db         660 AGAA---GCCTTAACCCAAGCCCCAGTGTTAAGCTTGCCAATGGGGCAAGACTTTTCTTT 716

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
Db         717 ATATGTCACAGAAAAAAACAGAAATAGCTCTAGGACTCCTTACACACGTCCAAGGGACCA 776

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAA 3554
Db         777 GCTCGCAACCCGTGGCATAGCTGAGTAAGGAAATTGATGTAGTGGAAAA 825

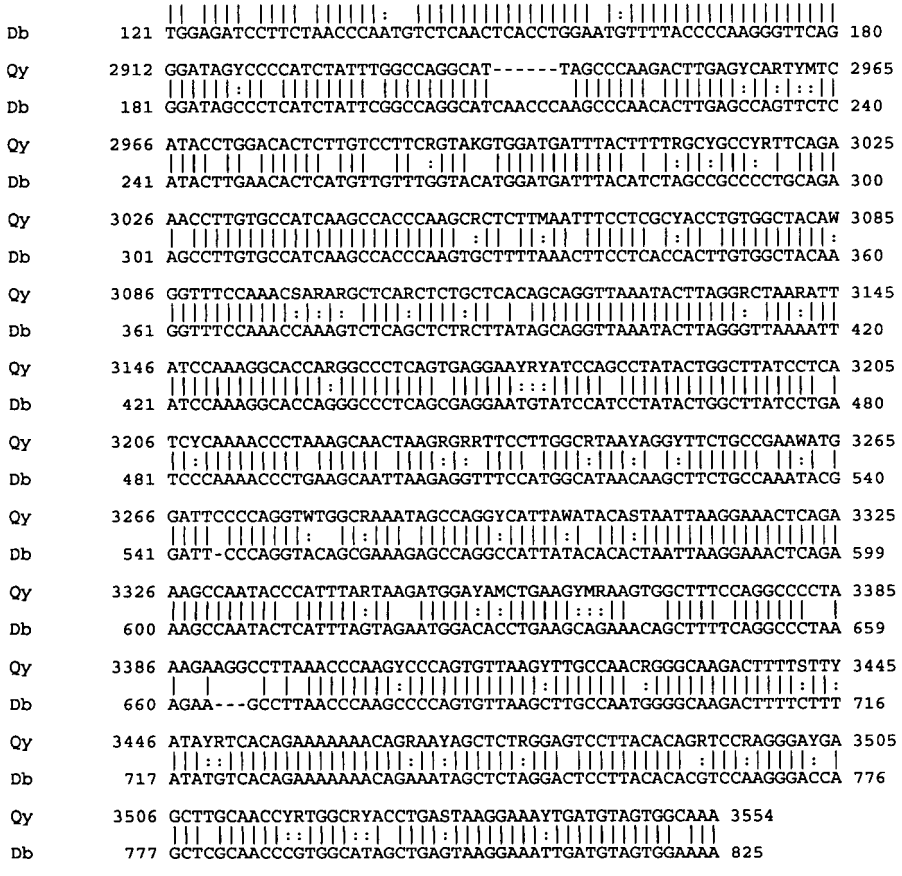

RESULT 85
US-09-925-065A-928581
; Sequence 928581, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928581
;   LENGTH: 825
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928581

  Query Match             8.5%;  Score 637.2;  DB 1;  Length 825;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  677;  Conservative   67;  Mismatches   75;  Indels   10;  Gaps    3;

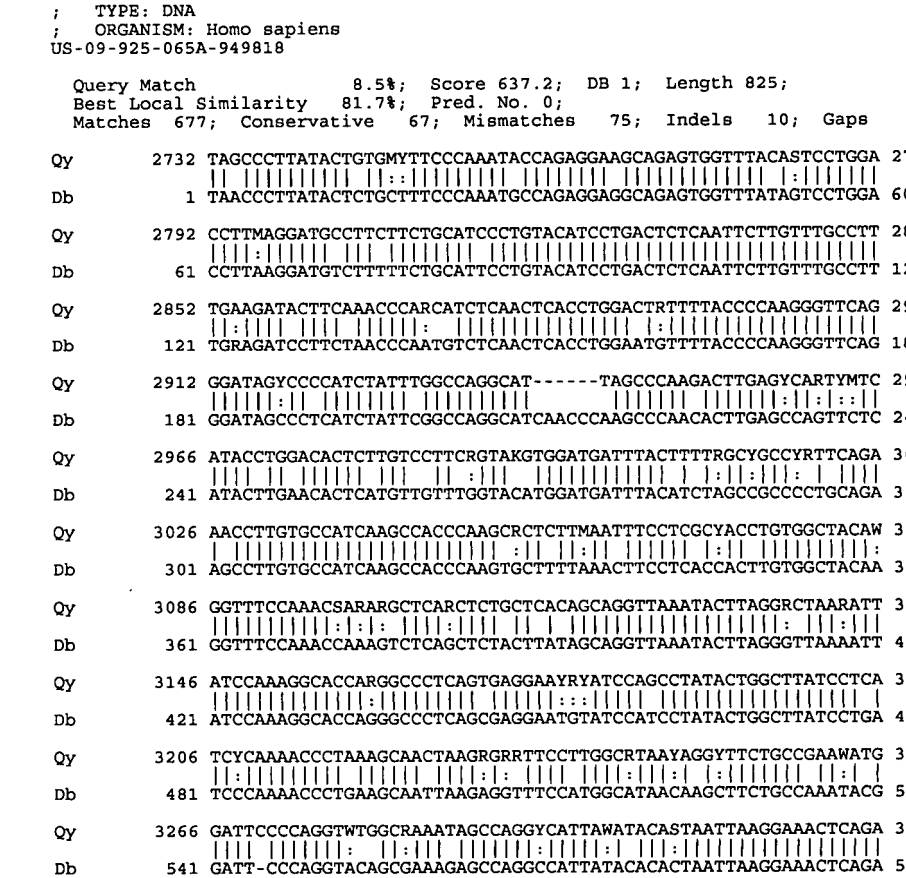
Qy       2732 TAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGA 2791
Db          1 TAACCCTTATACTCTGCTTTCCCAAATGCCAGAGGAGGCAGAGTGGTTTATAGTCCTGGA 60

Qy       2792 CCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 2851
Db         61 CCTTAAGGATGTCTTTTTCTGCATTCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 120

Qy       2852 TGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAG 2911
```



RESULT 86
US-09-925-065A-949818
; Sequence 949818, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 949818
; LENGTH: 825



RESULT 87
US-09-925-065A-949818
; Sequence 949818, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 949818
;   LENGTH: 825
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-949818

  Query Match             8.5%;  Score 637.2;  DB 1;  Length 825;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  677;  Conservative   67;  Mismatches   75;  Indels   10;  Gaps    3;

Qy       2732 TAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGA 2791
             || |||||||||| ||::|||||||| |||||||| |||||||||||| |:|||||||
Db          1 TAACCCTTATACTCTGCTTTCCCAAATGCCAGAGGAGGCAGAGTGGTTTATAGTCCTGGA 60

Qy       2792 CCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 2851
             ||||:|||||| ||| |||||||| ||||||||||||||||||||||||||||||||||
Db         61 CCTTAAGGATGTCTTTTTCTGCATTCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTT 120

Qy       2852 TGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAG 2911
             ||:|||| |||| |||||||:  |||||||||||||||| |:||||||||||||||||||
Db        121 TGRAGATCCTTCTAACCCAATGTCTCAACTCACCTGGAATGTTTTACCCCAAGGGTTCAG 180

Qy       2912 GGATAGYCCCCATCTATTTGGCCAGGCAT------TAGCCCAAGACTTGAGYCARTYMTC 2965
             ||||||:|| |||||||| ||||||||        ||||||| |||||||:||:|::||
Db        181 GGATAGCCCTCATCTATTCGGCCAGGCATCAACCCAAGCCCAACACTTGAGCCAGTTCTC 240

Qy       2966 ATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGA 3025
             |||| || |||||| |||  || :|||  |||||||||||| | |:||:|||: | ||||
Db        241 ATACTTGAACACTCATGTTGTTTGGTACATGGATGATTTACATCTAGCCGCCCCTGCAGA 300

Qy       3026 AACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             | |||||||||||||||||||||||| :|| ||:|| |||||| |:|| ||||||||||:
Db        301 AGCCTTGTGCCATCAAGCCACCCAAGTGCTTTTAAACTTCCTCACCACTTGTGGCTACAA 360

Qy       3086 GGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATT 3145
             ||||||||||:|:|: ||||:|||| || | ||||||||||||||||||: |||:|||
Db        361 GGTTTCCAAACCAAAGTCTCAGCTCTACTTATAGCAGGTTAAATACTTAGGGTTAAAATT 420

Qy       3146 ATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCA 3205
             |||||||||||||:|||||||| ||||||:::||||| |||||||||||||||||| |
Db        421 ATCCAAAGGCACCAGGGCCCTCAGCGAGGAATGTATCCATCCTATACTGGCTTATCCTGA 480

Qy       3206 TCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATG 3265
             ||:|||||||| |||||| ||||:|: |||| ||||:|||:| |:|||||| ||:| |
Db        481 TCCCAAAACCCTGAAGCAATTAAGAGGTTTCCATGGCATAACAAGCTTCTGCCAAATACG 540

Qy       3266 GATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGA 3325
             |||| |||||||:  ||:||| |||||||:|||||:| |||:|||||||||||||||||
Db        541 GATT-CCCAGGTACAGCGAAAGAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGA 599

Qy       3326 AAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
             |||||||||| |||||:|| |||||:|:||||||:::||   ||||| |||||||  |
Db        600 AAGCCAATACTCATTTAGTAGAATGGACACCTGAAGCAGAAACAGCTTTTCAGGCCCTAA 659

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
             | |    | | ||||||||:||||||||||||:||||||| :|||||||||||||:||:
Db        660 AGAA---GCCTTAACCCAAGCCCCAGTGTTAAGCTTGCCAATGGGGCAAGACTTTTCTTT 716

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
             |||::|||||||||||||||||:|||:|||||:||| |||||||||| :|||:|||||: |
Db        717 ATATGTCACAGAAAAAAACAGAAATAGCTCTAGGACTCCTTACACACGTCCAAGGGACCA 776

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAA 3554
             ||| ||||||::||||::| ||||:||||||||:|||||||||||| |||
Db        777 GCTCGCAACCCGTGGCATAGCTGAGTAAGGAAATTGATGTAGTGGAAAA 825


RESULT 88
US-09-925-065A-918078
; Sequence 918078, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918078
;   LENGTH: 736
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-918078

  Query Match             8.4%;  Score 629.3;  DB 1;  Length 736;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  641;  Conservative   56;  Mismatches   39;  Indels    3;  Gaps    2;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||| |||||||||| ||::||| |||||||||||||||||||| |||||||||:||||||
Db          1 CTAACCCTTATACTCTGCTTTCTCAAATACCAGAGGAAGCAGAATGGTTTACAGTCCTGG 60

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |  ||:|||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         61 ATATTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 120

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             |||||||| |||| ||  ||:|||||||||||||||||||| ||||||||||||||||||
Db        121 TTGAAGATCCTTCGAATGCAACATCTCAACTCACCTGGACTTTTTTACCCCAAGGGTTCA 180

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
             |||||||:|||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db        181 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACC 240

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||| ||||| :|||:|||||||||||||||||:||: ||::|||||||||||
Db        241 TGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACCT 300

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             |||||||||||||||||||:|||||:|||||||||||:|||||||||||||:|||||
Db        301 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTT 360

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||||||:|:|: | ||:|||||||||||||| |||||||||| ||:||||:||||||||
Db        361 CCAAACCAAAGACCCAGCTCTGCTCACAGCAGTTTAAATACTTGGGGCTAAAATTATCCA 420

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             |||||||||:|||||||||||||||:|:||| ||||||||||||||||||||||||:|
Db        421 AAGGCACCAGGGCCCTCAGTGAGGAACRTATCAAGCCTATACTGGCTTATCCTCATCCCC 480
```

```
Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db        481 AAATCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGTTTCTGCTGAATATGGATT- 539

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
Db        540 CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGCC 599

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
Db        600 AATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGA 657

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
Db        658 GGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTATG 717

Qy       3451 TCACAGAAAAAAACAGRAA 3469
Db        718 TCACAGAAAAAACAGGAAA 736


RESULT 89
US-09-925-065A-918078
; Sequence 918078, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918078
;   LENGTH: 736
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-918078

  Query Match             8.4%;  Score 629.3;  DB 1;  Length 736;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  641;  Conservative   56;  Mismatches   39;  Indels    3;  Gaps    2;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
Db          1 CTAACCCTTATACTCTGCTTTCTCAAATACCAGAGGAAGCAGAATGGTTTACAGTCCTGG 60

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
Db         61 ATATTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 120

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
Db        121 TTGAAGATCCTTCGAATGCAACATCTCAACTCACCTGGACTTTTTTACCCCAAGGGTTCA 180

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
Db        181 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACC 240

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
Db        241 TGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACCT 300

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
Db        301 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTT 360

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db        361 CCAAACCAAAGACCCAGCTCTGCTCACAGCAGTTTAAATACTTGGGGCTAAAATTATCCA 420

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
Db        421 AAGGCACCAGGGCCCTCAGTGAGGAACRTATCAAGCCTATACTGGCTTATCCTCATCCCC 480

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db        481 AAATCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGTTTCTGCTGAATATGGATT- 539

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
Db        540 CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGCC 599

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
Db        600 AATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAGA 657

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
Db        658 GGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTATG 717

Qy       3451 TCACAGAAAAAAACAGRAA 3469
Db        718 TCACAGAAAAAACAGGAAA 736


RESULT 90
US-09-925-065A-925162
; Sequence 925162, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925162
;   LENGTH: 791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925162

  Query Match             8.3%;  Score 623.6;  DB 1;  Length 791;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  717;  Conservative    3;  Mismatches   68;  Indels   10;  Gaps    5;

Qy       1651 AAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCA 1710
Db          1 ACACTTCAAAAGTCCACCTTAGGCCCGGAGCAAAACTTAGGAGCCCTATTGAACTTGGCA 60

Qy       1711 ACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAA 1770
```

```
                ||:|| |||||||||||||| |||||||||||||||| ||||  ||||||||||||||||
Db          61 ACCTCAGTTTTTTATAATAAAGATCAGGAGGAGCAGGTGGAATGGGACAAACGGGATTAA 120

Qy        1771 AAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAA 1830
                |||||||| ||| |||||||||||| |||||||||||||||||||||||||||||||||
Db         121 AAAAAAGGTCACTGCTTTAGTCATGGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGGAAA 180

Qy        1831 AGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGG 1890
                |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         181 AGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGG 240

Qy        1891 ACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATT 1950
                ||||||| | |:||||||| || ||||| |||||||||| | ||||||||||||||||
Db         241 ACACTTTTACARAGATTGTTTAAATAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTATG 300

Qy        1951 TCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGC 2010
                |||||||||||||||||||||||||||||||||||||||  || ||  | ||||||||||||
Db         301 TCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAAGTTCTCTGAGTCAGAAGC 360

Qy        2011 CACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGC 2070
                |||||||||||||||||||||||||||||||||||||||  |||||| ||| |||||||
Db         361 CACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGCAGCAAGCACCAGCCCATGC 420

Qy        2071 CATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCT 2127
                |||||||||||||||||| |||||||||||||||||||||||||||||| |   |||||
Db         421 CATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAAGTTAACTGTCT 480

Qy        2128 CCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAG 2187
                |||||||||||||| | ||||| |||||||||| || |||| ||||||||||||||||
Db         481 GCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTATCCCAGACAACTGTCCTCCAG 540

Qy        2188 ATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGC 2247
                |||||||||||| |   | |||  | || ||| | |||||||||||||||| :  ||
Db         541 ATCTGTCACTATCC---GAGGGGTCCTAGGACAGCCAGTCACTAGATACTTCT--CCCAG 595

Qy        2248 CACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAA 2307
                ||||||| | || |||||||| | ||| |||||||||||||||||||||||||| ||||
Db         596 CACTAAGCTGTG-ACTGGGGAACATTACTCTTTTCACATGCTTTTCTAATTATGCCTGAA 654

Qy        2308 AGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTG 2367
                | ||||||| |||||||||||||||||||| |||||||||||| ||||||||||||||||
Db         655 AACCCCACTCCCTTGTTAGGGAGAGACATTGTAGCAAAAGCAGGTGCCATTATACACCTG 714

Qy        2368 AACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGA 2427
                |||| |||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db         715 AACACAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCTTGA 773

Qy        2428 AGTCTGGGCAACAGAAGG 2445
                |||||||||||||||||
Db         774 AGTCTGGGCAACAGAAGG 791


RESULT 91
US-09-925-065A-925162
; Sequence 925162, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925162
;   LENGTH: 791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925162

  Query Match             8.3%;  Score 623.6;  DB 1;  Length 791;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  717;  Conservative    3;  Mismatches   68;  Indels   10;  Gaps    5;

Qy        1651 AAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCA 1710
                | |||||||||||  || |||||||||||||||||||| | ||||||||||||||||||
Db           1 ACACTTCAAAAGTCCACCTTAGGCCCGGAGCAAAACTTAGGAGCCCTATTGAACTTGGCA 60

Qy        1711 ACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAA 1770
                ||:|| |||||||||||| |||||||||||||||| |||| ||||||||||||||||||
Db          61 ACCTCAGTTTTTTATAATAAAGATCAGGAGGAGCAGGTGGAATGGGACAAACGGGATTAA 120

Qy        1771 AAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAA 1830
                |||||||| ||| |||||||||||| |||||||||||| |||||||||||||||||||||
Db         121 AAAAAAGGTCACTGCTTTAGTCATGGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGGAAA 180

Qy        1831 AGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGG 1890
                |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         181 AGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGG 240

Qy        1891 ACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATT 1950
                ||||||| | |:||||||| || ||||| |||||||||| | ||||||||||||||||
Db         241 ACACTTTTACARAGATTGTTTAAATAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTATG 300

Qy        1951 TCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGC 2010
                |||||||||||||||||||||||||||||||||||||||  || ||  | ||||||||||||
Db         301 TCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAAGTTCTCTGAGTCAGAAGC 360

Qy        2011 CACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGC 2070
                |||||||||||||||||||||||||||||||||||||||  |||||| ||| |||||||
Db         361 CACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGCAGCAAGCACCAGCCCATGC 420

Qy        2071 CATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCT 2127
                |||||||||||||||||| |||||||||||||||||||||||||||||| |   |||||
Db         421 CATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAAGTTAACTGTCT 480

Qy        2128 CCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAG 2187
                |||||||||||||| | ||||| |||||||||| || |||| ||||||||||||||||
Db         481 GCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTATCCCAGACAACTGTCCTCCAG 540

Qy        2188 ATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGC 2247
                |||||||||||| |   | |||  | || ||| | |||||||||||||||| :  ||
Db         541 ATCTGTCACTATCC---GAGGGGTCCTAGGACAGCCAGTCACTAGATACTTCT--CCCAG 595

Qy        2248 CACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAA 2307
                ||||||| | || |||||||| | ||| |||||||||||||||||||||||||| ||||
Db         596 CACTAAGCTGTG-ACTGGGGAACATTACTCTTTTCACATGCTTTTCTAATTATGCCTGAA 654

Qy        2308 AGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTG 2367
                | ||||||| |||||||||||||||||||| |||||||||||| ||||||||||||||||
Db         655 AACCCCACTCCCTTGTTAGGGAGAGACATTGTAGCAAAAGCAGGTGCCATTATACACCTG 714

Qy        2368 AACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGA 2427
                |||| |||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db         715 AACACAGGAGAAGGAACACCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCTTGA 773

Qy        2428 AGTCTGGGCAACAGAAGG 2445
                |||||||||||||||||
Db         774 AGTCTGGGCAACAGAAGG 791


RESULT 92
US-09-925-065A-941866
; Sequence 941866, Application US/09925065A
; Publication No. US20040181048A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941866
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-941866

  Query Match             8.3%;  Score 621.1;  DB 1;  Length 719;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  630;  Conservative   55;  Mismatches   34;  Indels    3;  Gaps    2;

Qy       2738 TTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMA 2797
             ||||||| ||::|||||||||||||||||||||||||||| |||||:|| ||||||||:|
Db          1 TTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGCTTACAGTCGTGGACCTTAA 60

Qy       2798 GGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGA 2857
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         61 GGATGCCTTTTTCTGCATCCCTGTACATCCTGACTGTCAATTCTTGTTTGCCTTTGAAGA 120

Qy       2858 TACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAG 2917
             | |||| ||||||:|||||||| |||||||||||:|||||||||||||||||| ||||||
Db        121 TCCTTCGAACCCAACATCTCAATTCACCTGGACTGTTTTACCCCAAGGGTTCAAGGATAG 180

Qy       2918 YCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACAC 2977
             :|||| | |||||||||||||||||||||||||||||:||:|::|||||||||||||||
Db        181 CCCCCTTTTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACAC 240

Qy       2978 TCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCA 3037
             | |||||||||:||| |||||||||||| ||||:||: ||::|||||||||||||||||
Db        241 TTTTGTCCTTCAGTACATGGATGATTTAATTTTAGCCACCTGTTCAGAAACCTTGTGCCA 300

Qy       3038 TCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACS 3097
             ||||||||||||||| :|||||:|||||||||||:|||||||||||||:|||||||||| :
Db        301 TCAAGCCACCCAAGAGCTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAATG 360

Qy       3098 ARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCAC 3157
             |:|:|||||:|||||||||||||||||||:||||||||:||||:|||||||||||||||
Db        361 AAAGGCTCAGCTCTGCTCACAGCAGGTTAAWTACTTAGGGCTAAAATTATCCAAAGGCAC 420

Qy       3158 CARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCT 3217
             ||:|||||||||||||||:::|||||||||| ||||||||||||||||:|||||||||||
Db        421 CAGGGCCCTCAGTGAGGAATGTATCCAGCCTACACTGGCTTATCCTCATCCCAAAACCCT 480

Qy       3218 AAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:| :|||||||||:|||:|||:|||||| |||:|||||| |||||||
Db        481 AAAGCAACTAAGAGCGTTCCTTGGCATAACAGGTTTCTGCGGAATATGGATT-CCCAGGT 539

Qy       3278 WTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCC 3337
             :||| :|||||||||| :|||||:|||||:||||||||||||||||||||||||||||||
Db        540 ATGGTGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCC 599

Qy       3338 ATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTA 3397
             |||||:|||||||||:|:|||| |:::|||||||||||||||||  ||  | | |||
Db        600 ATTTAGTAAGATGGACACCTGAGGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT-- 657

Qy       3398 AACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
             ||||||||:| ||||||||||:||||||| :|| ||||||||||:||:||| :|||||||
Db        658 AACCCAAGCCACAGTGTTAAGCTTGCCAAGGGGACAAGACTTTTCTTTATAAGTCACAGA 717

Qy       3458 AA 3459
             ||
Db        718 AA 719


RESULT 93
US-09-925-065A-941866
; Sequence 941866, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941866
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-941866

  Query Match             8.3%;  Score 621.1;  DB 1;  Length 719;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  630;  Conservative   55;  Mismatches   34;  Indels    3;  Gaps    2;

Qy       2738 TTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMA 2797
             ||||||| ||::|||||||||||||||||||||||||||| |||||:|| ||||||||:|
Db          1 TTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGCTTACAGTCGTGGACCTTAA 60

Qy       2798 GGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGA 2857
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         61 GGATGCCTTTTTCTGCATCCCTGTACATCCTGACTGTCAATTCTTGTTTGCCTTTGAAGA 120

Qy       2858 TACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAG 2917
             | |||| ||||||:|||||||| |||||||||||:|||||||||||||||||| ||||||
Db        121 TCCTTCGAACCCAACATCTCAATTCACCTGGACTGTTTTACCCCAAGGGTTCAAGGATAG 180

Qy       2918 YCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACAC 2977
             :|||| | |||||||||||||||||||||||||||||:||:|::|||||||||||||||
Db        181 CCCCCTTTTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACAC 240

Qy       2978 TCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCA 3037
             | |||||||||:||| |||||||||||| ||||:||: ||::|||||||||||||||||
Db        241 TTTTGTCCTTCAGTACATGGATGATTTAATTTTAGCCACCTGTTCAGAAACCTTGTGCCA 300

Qy       3038 TCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACS 3097
             ||||||||||||||| :|||||:|||||||||||:|||||||||||||:|||||||||| :
Db        301 TCAAGCCACCCAAGAGCTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAATG 360

Qy       3098 ARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCAC 3157
             |:|:|||||:|||||||||||||||||||:||||||||:||||:|||||||||||||||
Db        361 AAAGGCTCAGCTCTGCTCACAGCAGGTTAAWTACTTAGGGCTAAAATTATCCAAAGGCAC 420

Qy       3158 CARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCT 3217
             ||:|||||||||||||||:::|||||||||| ||||||||||||||||:|||||||||||
Db        421 CAGGGCCCTCAGTGAGGAATGTATCCAGCCTACACTGGCTTATCCTCATCCCAAAACCCT 480
```

```
Qy       3218 AAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
Db        481 AAAGCAACTAAGAGCGTTCCTTGGCATAACAGGTTTCTGCGGAATATGGATT-CCCAGGT 539

Qy       3278 WTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCC 3337
Db        540 ATGGTGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCC 599

Qy       3338 ATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTA 3397
Db        600 ATTTAGTAAGATGGACACCTGAGGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT-- 657

Qy       3398 AACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
Db        658 AACCCAAGCCACAGTGTTAAGCTTGCCAAGGGGACAAGACTTTTCTTTATAAGTCACAGA 717

Qy       3458 AA 3459
Db        718 AA 719
```

RESULT 94
US-09-925-065A-936448/c
; Sequence 936448, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936448
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-936448

Query Match 8.3%; Score 619.1; DB 1; Length 762;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 673; Conservative 54; Mismatches 35; Indels 11; Gaps 6;

```
Qy       3244 TAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWAT 3303
Db         762 TAACAGGTTTCTGCCGAATATGGATT-CCTAGGTATGGCAAAATAGCCAGACCATTATAT 704

Qy       3304 ACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYM 3363
Db         703 ACACTAATTAAGGAAACTCAGAAATCCAATACCCATTTAGTAAGATGGACACCTGAATCA 644

Qy       3364 RAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGC 3423
Db         643 GAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGT 586

Qy       3424 CAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTC 3483
Db         585 CAATGGGGCAAGACTTGACTTTATATGTCACAGAAAAGAACAGGAATAGCTCTAGGAGTC 526

Qy       3484 CTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGAT 3543
Db         525 CTTACACAGGTCCGAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGAT 466

Qy       3544 GTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTA 3603
Db         465 GTAGTGGC-AAGGGTTGGCCTCATGGTTTACAGGTAGTGGCAGCAGTAGCAGGCTTAGTA 407

Qy       3604 TCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTG 3663
Db         406 TCTGAAGCAGTTAAAATAATACAGG-----GATCTTACTGTGTGGACATC-CATGATGTG 353

Qy       3664 AAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRT 3723
Db         352 AATGGCATACTAACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTACTTAAATAT 293

Qy       3724 CAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCA 3783
Db         292 CAGGCTCTATTACTTGAAGGGCCAGTGCTGCRACTGCGCACTTGTGCAACTCTTAACCCA 233

Qy       3784 GYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCT 3843
Db         232 GCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCT 173

Qy       3844 CAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAA 3903
Db         172 CAAACATACGCCACTTGAGGGGACATTTTAGAGGTTCCCTTGACTGATCCCGACCTT-AA 114

Qy       3904 CTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGT 3963
Db         113 CTTGTGTACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGCGAGGTATGCAGT 54

Qy       3964 GGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db          53 GGTCAGTGATAATGGAATACTTGAAAGTAATCCCTTCACTGCAGGAACTAGTG 1
```

RESULT 95
US-09-925-065A-936448/c
; Sequence 936448, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936448
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-936448

Query Match 8.3%; Score 619.1; DB 1; Length 762;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 673; Conservative 54; Mismatches 35; Indels 11; Gaps 6;

```
Qy       3244 TAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWAT 3303
Db         762 TAACAGGTTTCTGCCGAATATGGATT-CCTAGGTATGGCAAAATAGCCAGACCATTATAT 704

Qy       3304 ACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYM 3363
```

```
                |||:|||||||||||||||||||| ||||||||||||||:|||||||||:|:|||||  ::
Db        703 ACACTAATTAAGGAAACTCAGAAATCCAATACCCATTTAGTAAGATGGACACCTGAATCA 644

Qy       3364 RAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGC 3423
              :|||||||||||||||||  ||  | | |||  ||||||||:|||||||||||||:|||
Db        643 GAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGT 586

Qy       3424 CAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTC 3483
              ||| :||||||||||||  :||:|||::||||||||| |||||:||:|||||:||||||
Db        585 CAATGGGGCAAGACTTGACTTTATATGTCACAGAAAAGAACAGGAATAGCTCTAGGAGTC 526

Qy       3484 CTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGAT 3543
              |||||||||:|||:|||||: |||||||||||::||||::||||||:||||||||:||||
Db        525 CTTACACAGGTCCGAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGAT 466

Qy       3544 GTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTA 3603
              |||||||| ||||||||:|:|||| |||||: |||||||| |||||||||| ::|||||
Db        465 GTAGTGGC-AAGGGTTGGCCTCATGGTTTACAGGTAGTGGCAGCAGTAGCAGGCTTAGTA 407

Qy       3604 TCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTG 3663
              |||||||||||||||||||||||||      |||||||||||||||||||| |||||:|||
Db        406 TCTGAAGCAGTTAAAATAATACAGG-----GATCTTACTGTGTGGACATC-CATGATGTG 353

Qy       3664 AAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRT 3723
              ||::|||||||| |||||||||||||||||||||||||||||: ||||||||||:|
Db        352 AATGGCATACTAACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTACTTAAATAT 293

Qy       3724 CAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCA 3783
              |||||||||||||||||:|||||||||||||||| |||||||||||||||||||||||||
Db        292 CAGGCTCTATTACTTGAAGGGCCAGTGCTGCRACTGCGCACTTGTGCAACTCTTAACCCA 233

Qy       3784 GYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCT 3843
              |:| |||||||||||||||||||||||||||||:||:|||||||||||||||:||||| ||
Db        232 GCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCT 173

Qy       3844 CAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAA 3903
              ||||| || |||||| |||||||| ||:|||||:|||||:||||||||||:|||||| ||
Db        172 CAAACATACGCCACTTGAGGGGACATTTTAGAGGTTCCCTTGACTGATCCCGACCTT-AA 114

Qy       3904 CTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGT 3963
              ||||| |||||||||||||||||||||||||||||||||||||||| |||||:| ||||||||||
Db        113 CTTGTGTACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGCGAGGTATGCAGT 54

Qy       3964 GGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
              |||||||||||||||||:|||||||||||||| ||||| ||||||||||||
Db          53 GGTCAGTGATAATGGAATACTTGAAAGTAATCCCTTCACTGCAGGAACTAGTG 1


RESULT 96
US-09-925-065A-937950/c
; Sequence 937950, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937950
;   LENGTH: 729
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937950

  Query Match             7.9%;  Score 593.7;  DB 1;  Length 729;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  624;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
              |||| |||||:|||| |||||||||||||||||||||||||||||||||:||:|: |||
Db        729 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 670

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
              || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        669 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 610

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
              |||||||| |||||||||||||||:|||||:||||||||||:|||||||||
Db        609 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 556

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                 ||||||| |:|:|||||:|||||||||||||||||||||||||||:||||:||||
Db        555 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 499

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
              ||||||| |||||:||||||||| ||||||:::||||||| ||||||||||||||||||
Db        498 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 439

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
              |:|||||||||||||||||||:|::|||||||:|||:|| :||||||||:||||
Db        438 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 379

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
              ||| ||||| |:  | :|||||||||||:|||||: |||| ||||||||||||||||||
Db        378 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 320

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
              |||||||||||| ||:|||||||||:|:|||||:::|||||||||||||||||  ||
Db        319 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 260

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :|||||| ||||:||||||||:|||||||||||||:||:|
Db        259 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 202

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
              ||::|||||| || |||||||:||:|||||||:|||||||||| |||:|| :|||||:|||
Db        201 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 143

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              |||||||||::||| ::|||||||:|||||||||||||||||||||||||:|:|||
Db        142 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 83

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
              |||||||:||||| |||| ||||||||||||::|||||||||||||||||||||||||||
Db         82 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 23

Qy       3627 GGGRAGAGATCTTACTGTGTGG 3648
              |||:||||||||| ||||||||
Db         22 GGGAAGAGATCTTTCTGTGTGG 1


RESULT 97
US-09-925-065A-937950/c
; Sequence 937950, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937950
;   LENGTH: 729
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937950

  Query Match             7.9%;  Score 593.7;  DB 1;  Length 729;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  624;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
             |||| ||||||:|||| |||||||||||||||||||||||||||||||||:||:|: |||
Db        729 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 670

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
             || ||||||| |||||| |||||:|||:| |||||||||| || |:||: ||::|||||||
Db        669 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 610

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
             |||||||| |||||||||||||||||:|||||:||||||||||:|||||||||
Db        609 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 556

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                |||||||  |:|:|||||:|||||||||||||||||||||||||||||||:||||:||||
Db        555 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 499

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||| |||||:|||||||||| ||||||:::|||||||  ||||||||||||||||||||
Db        498 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 439

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
             |:|||||||||||||||||||||:|::|||||||||:|||:|| :||||||||||:||||
Db        438 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 379

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
             ||| ||||| |:  | :|||||||||||:|||||: |||| ||||||||||||||||||||
Db        378 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 320

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
             |||||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||||  ||
Db        319 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 260

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :|||||  ||||:||||||||:|||||||||||||:||:|
Db        259 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 202

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| || |||||:||:|||||:|||||||||| |||:|| :|||||:|||
Db        201 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 143

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||| ::||||||:||||||||:|||||||||||||||||||||:|:|||
Db        142 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 83

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:||||| |||| |||||||||||::|||||||||||||||||||||||||||
Db         82 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 23

Qy       3627 GGGRAGAGATCTTACTGTGTGG 3648
             |||:|||||||||| ||||||||
Db         22 GGGAAGAGATCTTTCTGTGTGG 1

RESULT 98
US-09-925-065A-937711/c
; Sequence 937711, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937711
;   LENGTH: 727
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937711

  Query Match             7.9%;  Score 591.7;  DB 1;  Length 727;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  622;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
             |||| ||||||:|||| |||||||||||||||||||||||||||||||||:||:|: |||
Db        727 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 668

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
             || ||||||| |||||| |||||:|||:| |||||||||| || |:||: ||::|||||||
Db        667 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 608

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
             |||||||| |||||||||||||||||:|||||:||||||||||:|||||||||
Db        607 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 554

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                |||||||  |:|:|||||:|||||||||||||||||||||||||||||||:||||:||||
Db        553 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 497

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||| |||||:|||||||||| ||||||:::|||||||  ||||||||||||||||||||
Db        496 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 437

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
             |:|||||||||||||||||||||:|::|||||||||:|||:|| :||||||||||:||||
Db        436 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 377

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
             ||| ||||| |:  | :|||||||||||:|||||: |||| ||||||||||||||||||||
Db        376 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 318

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
             |||||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||||  ||
Db        317 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 258

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :|||||  ||||:||||||||:|||||||||||||:||:|
Db        257 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 200

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| || |||||:||:|||||:|||||||||| |||:|| :|||||:|||
Db        199 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 141
```

```
Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||| ::|||||:||||||||:|||||||||||||||||||||:|:|||
Db        140 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 81

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:||||| |||| ||||||||||::||||||||||||||||||||||||||||
Db         80 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 21

Qy       3627 GGGRAGAGATCTTACTGTGT 3646
             |||:|||||||||| ||||||
Db         20 GGGAAGAGATCTTTCTGTGT 1


RESULT 99
US-09-925-065A-937711/c
; Sequence 937711, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937711
;   LENGTH: 727
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937711

  Query Match             7.9%;  Score 591.7;  DB 1;  Length 727;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  622;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
             |||| |||||||:|||| |||||||||||||||||||||||||||||||:||:|: |||
Db        727 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 668

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
             || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        667 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 608

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
             |||||||| ||||||||||||||||:|||||:||||||||||:|||||||||
Db        607 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 554

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                |||||||  |:|:|||||:|||||||||||||||||||||||||||||:||||:||||
Db        553 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 497

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||| |||||:|||||||||| ||||||:::|||||||| ||||||||||||||||||
Db        496 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 437

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
             |:|||||||||||||||||||:|::|||||||||:|||:|| :||||||||||:||||
Db        436 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 377

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
             ||| ||||| |:  | :||||||||||||:|||||: |||| ||||||||||||||||||
```

```
Db        376 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 318

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
             |||||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||||  ||
Db        317 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 258

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :||||||  ||||:||||||||:|||||||||||||:||:|
Db        257 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 200

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| || ||||||:||:||||||:|||||||||| |||:|| :|||||:|||
Db        199 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 141

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||| ::|||||:||||||||:|||||||||||||||||||||:|:|||
Db        140 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 81

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:||||| |||| ||||||||||::||||||||||||||||||||||||||||
Db         80 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 21

Qy       3627 GGGRAGAGATCTTACTGTGT 3646
             |||:|||||||||| ||||||
Db         20 GGGAAGAGATCTTTCTGTGT 1


RESULT 100
US-09-925-065A-954498
; Sequence 954498, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 954498
;   LENGTH: 745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-954498

  Query Match             7.9%;  Score 590.8;  DB 1;  Length 745;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  654;  Conservative   54;  Mismatches   37;  Indels   28;  Gaps    6;

Qy       3239 TGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
             ||||:|||:|||:|||||||||:|  |||| |||||||:  ||:|||||||||| :|||
Db          1 TGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAGCAAAATAGCCAGACCAT 59

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||| |||||:|||||||:|:||
Db         60 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT- 118

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             || :::|||||||||||||||||  ||  | | |||  ||||||||:||||||||| ||
Db        119 AAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAG 176

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
```

```
                :|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||: |||||:|
Db        177 CTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATCGCTCTAG 235

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||||||:|||:|||||:|||||||||||::||||::|||||| ||||||||
Db        236 GAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCCATGGCATACCTGAATAAGGAAA 295

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
             :|||||||||||||||||||:|:|||||||||:|||||||| |||||||||||| :
Db        296 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCGGCAGTAGCAGTAT 355

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
             |||| |||||||||||||||||||||:||||||||||||| ||||||||||||
Db        356 TAGTGTCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTATGGACATCTCATG 415

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
             | |||||:: |||||||||||||||||| ||||||||||||||||||||:|||||||||
Db        416 ACGTGAATGACATACTCACTGCTAAAGAAGACTTGTGGCTGTCAGACAACTGTTTACTTA 475

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTA 3778
             | |:||||||||||||||||||:|||||||                        ||| |||
Db        476 ACTATCAGGCTCTATTACTTGAAGGGCCAGT----------------------ACTGTTA 513

Qy       3779 ACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAA 3838
             ||||||:| |||||||||||||||||||||||||||||::|:|||||||||||||:|||
Db        514 ACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAA 573

Qy       3839 TTTCTCAAACCTATGCCACTCGAGGGGACCTTYTATGARGTTCCYTTGACTGATCCYGACC 3898
             || ||||||||| ||| |||||||||||||:||||:|||||:|||||||||||::|||
Db        574 TTCCTCAAACCTGTGCTGCTCGAGGGGACCTTTTAGAGGTTCCCTTGACTGATCCCRACC 633

Qy       3899 TTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT 3958
             |||||||||||||||||||||||||| ||||||||||||||||||||||:||| |||
Db        634 -TCAACTTGTATACTGATGGAAGTTCCCTTGTAGAAAAAGGACTTCGAAAAGTGGGCTAT 692

Qy       3959 GCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAAC 4011
             || |||||||||||||  |||||:||||||||||||||||||||||||||||
Db        693 GCGGTGGTCAGTGATAACAGAATACTTGAAAGTAATCCCCTCACTCCAGGAAC 745


RESULT 101
US-09-925-065A-954498
; Sequence 954498, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 954498
;   LENGTH: 745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-954498

  Query Match             7.9%;  Score 590.8;  DB 1;  Length 745;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  654;  Conservative   54;  Mismatches   37;  Indels   28;  Gaps    6;

Qy       3239 TGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
             ||||:|||:|||:||||||||||:|  |||| |||||||:  ||:|||||||||| :|||
Db          1 TGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAGCAAAATAGCCAGACCAT 59

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||| |||||:|||||||||:|:||
Db         60 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT- 118

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             || :::|||||||||||||||||||  ||  | | |||  |||||||||:|||||||||| ||
Db        119 AAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAG 176

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||: |||||:|
Db        177 CTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATCGCTCTAG 235

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||||||:|||:||||||:|||||||||||::||||::|||||| ||||||||
Db        236 GAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCCATGGCATACCTGAATAAGGAAA 295

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
             :|||||||||||||||||||:|:|||||||||:|||||||| |||||||||||| :
Db        296 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCGGCAGTAGCAGTAT 355

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
             |||| |||||||||||||||||||||:||||||||||||| ||||||||||||
Db        356 TAGTGTCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTATGGACATCTCATG 415

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
             | |||||:: |||||||||||||||||| ||||||||||||||||||||:|||||||||
Db        416 ACGTGAATGACATACTCACTGCTAAAGAAGACTTGTGGCTGTCAGACAACTGTTTACTTA 475

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTA 3778
             | |:||||||||||||||||||:|||||||                        ||| |||
Db        476 ACTATCAGGCTCTATTACTTGAAGGGCCAGT----------------------ACTGTTA 513

Qy       3779 ACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAA 3838
             ||||||:| |||||||||||||||||||||||||||||::|:|||||||||||||:|||
Db        514 ACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAA 573

Qy       3839 TTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACC 3898
             || ||||||||| ||| |||||||||||||:||||:|||||:|||||||||||::|||
Db        574 TTCCTCAAACCTGTGCTGCTCGAGGGGACCTTTTAGAGGTTCCCTTGACTGATCCCRACC 633

Qy       3899 TTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT 3958
             |||||||||||||||||||||||||| ||||||||||||||||||||||:||| |||
Db        634 -TCAACTTGTATACTGATGGAAGTTCCCTTGTAGAAAAAGGACTTCGAAAAGTGGGCTAT 692

Qy       3959 GCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAAC 4011
             || |||||||||||||  |||||:||||||||||||||||||||||||||||
Db        693 GCGGTGGTCAGTGATAACAGAATACTTGAAAGTAATCCCCTCACTCCAGGAAC 745


RESULT 102
US-09-925-065A-926262/c
; Sequence 926262, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926262
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926262

  Query Match             7.9%;  Score 589.7;  DB 1;  Length 725;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  620;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
             |||| |||||||:|||| |||||||||||||||||||||||||||||||||:||:|: |||
Db        725 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 666

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
             || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        665 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 606

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
             |||||||| |||||||||||||||||:|||||:||||||||||:|||||||||
Db        605 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 552

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                 ||||||| |:|:|||||:|||||||||||||||||||||||||||||:||||:||||
Db        551 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 495

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||| |||||:||||||||| ||||||:::||||||| ||||||||||||||||||||
Db        494 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 435

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
             |:||||||||||||||||||||:|::||||||||:|||:|| :||||||||||:||||
Db        434 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 375

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
             ||| ||||| |:  | :||||||||||:|||||: |||| ||||||||||||||||||||
Db        374 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 316

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
             |||||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||||  ||
Db        315 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 256

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :||||||  ||||:|||||||:|||||||||||||:||:|
Db        255 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 198

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| || ||||||:||:||||||:|||||||||| |||:|| :|||||:|||
Db        197 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 139

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||| ::||||||:||||||||:|||||||||||||||||||:|:|||
Db        138 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 79

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:||||| |||| ||||||||||::|||||||||||||||||||||||||||||
Db         78 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 19

Qy       3627 GGGRAGAGATCTTACTGT 3644
              |||:|||||||||| ||||
Db         18 GGGAAGAGATCTTTCTGT 1


RESULT 103
US-09-925-065A-926262/c
; Sequence 926262, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926262
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926262

  Query Match             7.9%;  Score 589.7;  DB 1;  Length 725;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  620;  Conservative   66;  Mismatches   39;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
             |||| |||||||:|||| |||||||||||||||||||||||||||||||||:||:|: |||
Db        725 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 666

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
             || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        665 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 606

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
             |||||||| |||||||||||||||||:|||||:||||||||||:|||||||||
Db        605 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 552

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
                 ||||||| |:|:|||||:|||||||||||||||||||||||||||||:||||:||||
Db        551 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 495

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
             ||||||| |||||:||||||||| ||||||:::||||||| ||||||||||||||||||||
Db        494 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 435

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
             |:||||||||||||||||||||:|::||||||||:|||:|| :||||||||||:||||
Db        434 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 375

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
             ||| ||||| |:  | :||||||||||:|||||: |||| ||||||||||||||||||||
Db        374 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 316

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
             |||||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||||  ||
Db        315 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 256

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :||||||  ||||:|||||||:|||||||||||||:||:|
Db        255 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 198

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| || ||||||:||:||||||:|||||||||| |||:|| :|||||:|||
Db        197 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 139

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||| ::||||||:||||||||:|||||||||||||||||||:|:|||
Db        138 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 79

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:||||| |||| ||||||||||::|||||||||||||||||||||||||||||
Db         78 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 19
```

```
Qy       3627 GGGRAGAGATCTTACTGT 3644
              |||:|||||||||| ||||
Db         18 GGGAAGAGATCTTTCTGT 1


RESULT 104
US-09-925-065A-937967/c
; Sequence 937967, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937967
;   LENGTH: 720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937967

  Query Match             7.8%;  Score 586.3;  DB 1;  Length 720;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  616;  Conservative   66;  Mismatches   38;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
              |||| ||||||:|||| |||||||||||||||||||||||||||||||:||:|: |||
Db        720 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 661

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
              || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        660 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 601

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
              |||||||| ||||||||||||||||||:|||||:||||||||||:|||||||||
Db        600 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 547

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
              ||||||| |:|:|||||:|||||||||||||||||||||||||||||:||||:|||||
Db        546 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 490

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
              ||||||| |||||:||||||||| ||||||:::||||||| |||||||||||||||||||
Db        489 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 430

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
              |:||||||||||||||||||||:|::|||||||||:|||:|| :||||||||||:||||
Db        429 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 370

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
              ||| ||||| |:  | :|||||||||:|||||: |||| ||||||||||||||||||||
Db        369 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 311

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
              |||||||||||| ||:|||||||||:|:|||||::::|||||||||||||||  ||
Db        310 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 251

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | | |||  ||||||| :||||||  ||||:||||||||:||||||||||||||:||:|

Db        250 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 193

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
              ||::|||||| || |||||||:||:|||||||:|||||||||||| |||:|| :|||||:|||
Db        192 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 134

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              |||||||||::||| ::|||||||:||||||||:|||||||||||||||||||:|:|||
Db        133 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 74

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
              |||||||:||||| |||| ||||||||||::|||||||||||||||||||||||||||
Db         73 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 14

Qy       3627 GGGRAGAGATCTT 3639
              |||:|||||||||
Db         13 GGGAAGAGATCTT 1


RESULT 105
US-09-925-065A-937967/c
; Sequence 937967, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937967
;   LENGTH: 720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-937967

  Query Match             7.8%;  Score 586.3;  DB 1;  Length 720;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  616;  Conservative   66;  Mismatches   38;  Indels   13;  Gaps    4;

Qy       2907 TTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCA 2966
              |||| ||||||:|||| |||||||||||||||||||||||||||||||:||:|: |||
Db        720 TTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTGTCA 661

Qy       2967 TACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAA 3026
              || ||||||| |||||| ||||:|||:| |||||||||| || |:||: ||::|||||||
Db        660 TAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTTAATTATAGCCACCCATTCAGAA 601

Qy       3027 ACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWG 3086
              |||||||| ||||||||||||||||||:|||||:||||||||||:|||||||||
Db        600 ACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCGCTACCTGTGGC------ 547

Qy       3087 GTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA 3146
              ||||||| |:|:|||||:|||||||||||||||||||||||||||||:||||:|||||
Db        546 ---TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTA 490

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
              ||||||| |||||:||||||||| ||||||:::||||||| |||||||||||||||||||
Db        489 TCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGTATACTGGCTTATCCTCAT 430

Qy       3207 CYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
```

```
Db        429 CCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGCCTTCTGCCGAATATGG 370

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326

Db        369 ATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTACATTAATTAAGGAAACTCAGAA 311

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386

Db        310 AGCCAATACCCATATAGTAAGATGGACACCTGAARCAGAAGTGGCTTTCCAGGCCCTAAA 251

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446

Db        250 GAAGGCCCT--AACCCAAACCCCAGTTGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTA 193

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506

Db        192 TATGTCACAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTGAGGGATGAG 134

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566

Db        133 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCA 74

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626

Db          73 TTGTTTATGGGTAATGGTAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACA 14

Qy       3627 GGGRAGAGATCTT 3639

Db          13 GGGAAGAGATCTT 1


RESULT 106
US-08-979-847-9
; Sequence 9, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098

Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158

Db          61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218

Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277

Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336

Db        241 ACAACCAAAATAGCCAGACCATTATATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396

Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456

Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516

Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576

Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636

Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696

Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756

Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782

Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 107
US-10-430-442-9
; Sequence 9, Application US/10430442
; Publication No. US20030186391A1
;  GENERAL INFORMATION:
;       APPLICANT: Herve PERRON
;                  Frederic BESEME
```

```
;                  Frederic BEDIN
;                  Glaucia PARANHOS-BACCALA
;                  Florence KOMURIAN-PRADEL
;                  Colette JOLIVET
;                  Bernard MANDRAND
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 92
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Oliff & Berridge
;            STREET: 700 South Washington Street, Suite 300
;            CITY: Alexandria
;            STATE: Virginia
;            COUNTRY: U.S.A.
;            ZIP: 22314
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/430,442
;            FILING DATE: 07-May-2003
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/691,563
;            FILING DATE: 02-AUG-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Berridge, William P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 38588
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 9:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 741 base pairs
;            TYPE: nucleotide
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-430-442-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:|||||||| |:|||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              |||||||||||:|::|||||| ||:| | :|||:|||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:||||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
               |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              |||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: |||||||| |:|||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 108
US-10-114-104-9
; Sequence 9, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/114,104
;            FILING DATE: 03-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/979,847
;            FILING DATE: 26-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 9:
```

```
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 741 base pairs
;            TYPE: nucleotide
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-114-104-9

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             ||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db             1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db            61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db           121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:|||||||||:|      | ||||||| 
Db           181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
             :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||| 
Db           241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||   ||  | | |||
Db           301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: |||||
Db           357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||: 
Db           416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::||||||  ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db           475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||||:||||||
Db           535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db           595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: |||||||| |:|||||||||||||||||: |||||||||| :|
Db           655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db           715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 109
US-09-925-065A-936449/c
; Sequence 936449, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936449
;   LENGTH: 709
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936449

  Query Match             7.8%;  Score 585.9;  DB 1;  Length 709;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  619;  Conservative   54;  Mismatches   35;  Indels    9;  Gaps    4;

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||| |||||||||||||||||||:||||||||||:|||||| || 
Db           708 TGCTTTCCCAAATACCGGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGTCT 649

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             | |||||||||||||||||||||||||| |||||||||| ||||||||||||   |||||| 
Db           648 TTTTCTGCATCCCTGTACATCCTGACTATCAATTCTTGGTTGCCTTTGAAGAGCCTTCAA 589

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:| |||||||||||||||||:|||||||||||||||||||| |||||:|||||| 
Db           588 ACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGAGATAGCCCCCATC 529

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             ||||     |||||||||||||||||||:||:|::|||||:|||||||||||||||||| 
Db           528 TATT-----AGGCATTAGCCCAAGACTTGAGTCAATTCTCATAYCTGGACACTCTTGTCC 474

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             || :||| |||||||||||||||:||:|||: |||||||||||||||||||||||||||| 
Db           473 TTTGGTACATGGATGATTTACTTTTAGCTGCCCTTTCAGAAACCTTGTGCCATCAAGCCA 414

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||| :|||||: |||||||||||:|||||||||||||:|||||||||| |:|:|||| 
Db           413 CCCAAGAGCTCTTACATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTC 354

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||| |||||||||||||:|||||:|| |||||||||||||:|||||| 
Db           353 AGCTCTGCTCACAGCAGTTTAAATACTTAGGGCTAAAATCATCCAAAGGCACCAGGGCCC 294

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||| |||||:::|||| |||||||||||||||||||||||||:|||||||||||||| 
Db           293 TCAGTCAGGAATGTATCCTGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 234

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:| :|||||||||:|||:|||:|||||||||:|||||| || ||||:||||:|| 
Db           233 TAAGAG-GTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-CCTAGGTATGGCAAA 176

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||| :|||||:|||||:|||||||||||||||||| ||||||||||||:|| 
Db           175 ATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAATCCAATACCCATTTAGTA 116

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||| :::|||||||||||||||||  ||  | | |||  |||||||| 
Db           115 AGATGGACACCTGAATCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAG 58

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAA 3462
             :|||||||||||:||| ||| :|||||||||||  :||:|||::|||||||||| | 
Db            57 CCCCAGTGTTAAGCTTGTCAATGGGGCAAGACTTGACTTTATATGTCACAGAAAAGA 1
```

```
RESULT 110
US-09-925-065A-936449/c
; Sequence 936449, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936449
;   LENGTH: 709
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936449

  Query Match             7.8%;  Score 585.9;  DB 1;  Length 709;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  619;  Conservative   54;  Mismatches   35;  Indels    9;  Gaps    4;

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||| |||||||||||||||||||:||||||||||:|||||| ||
Db        708 TGCTTTCCCAAATACCGGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGTCT 649

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             | |||||||||||||||||||||||| |||||||||| |||||||||||||  ||||||
Db        648 TTTTCTGCATCCCTGTACATCCTGACTATCAATTCTTGGTTGCCTTTGAAGAGCCTTCAA 589

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:| |||||||||||||||||:|||||||||||||||||| |||||:|||||||
Db        588 ACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGAGATAGCCCCCATC 529

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             ||||     |||||||||||||||||||||:||:|::|||||:||||||||||||||||
Db        528 TATT-----AGGCATTAGCCCAAGACTTGAGTCAATTCTCATAYCTGGACACTCTTGTCC 474

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             || :|||  |||||||||||||||||:||:|||: |||||||||||||||||||||||||
Db        473 TTTGGTACATGGATGATTTACTTTTAGCTGCCCTTTCAGAAACCTTGTGCCATCAAGCCA 414

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||| :|||||: |||||||||:|||||||||||||:|||||||||| |:|:||||
Db        413 CCCAAGAGCTCTTACATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTC 354

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||| |||||||||||||:||||:|| |||||||||||||:|||||
Db        353 AGCTCTGCTCACAGCAGTTTAAATACTTAGGGCTAAAATCATCCAAAGGCACCAGGGCCC 294

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||| |||||:::|||| |||||||||||||||||||||:|||||||||||||||||
Db        293 TCAGTCAGGAATGTATCCTGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 234

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:| :|||||||||:|||:|||:||||||||||:||||||| || ||||:||||:||
Db        233 TAAGAG-GTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-CCTAGGTATGGCAAA 176

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||| :|||||:|||||:|||||||||||||||||||| ||||||||||||||:||
```

```
Db        175 ATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAATCCAATACCCATTTAGTA 116

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||| :::|||||||||||||||||||  ||  | | |||  ||||||||
Db        115 AGATGGACACCTGAATCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAG 58

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAA 3462
             :||||||||||||:||| ||| :|||||||||||  :||:|||::|||||||||| |
Db         57 CCCCAGTGTTAAGCTTGTCAATGGGGCAAGACTTGACTTTATATGTCACAGAAAAGA 1


RESULT 111
US-10-717-580-47
; Sequence 47, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-47

  Query Match             7.8%;  Score 582.1;  DB 1;  Length 600;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  586;  Conservative   11;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420
```

```
Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              |||||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 M 601
              :
Db        600 A 600
```

RESULT 112
US-09-925-065A-937161/c
; Sequence 937161, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937161
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-937161

Query Match 7.8%; Score 581.5; DB 1; Length 720;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 621; Conservative 51; Mismatches 47; Indels 5; Gaps 4;

```
Qy       3289 GCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGA 3348
              | ||| :|||||:|||||:|||||||||||||||||||| ||||| |||||||:|||||
Db        720 GACAGACCATTATATATACACTAATTAAGGAAACTCAGAAAGTCAATATCCATTTAGTAAGA 661

Qy       3349 TGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
              ||||:|:||||| :::|||||||||||||||||    ||  | | ||  |||||||| |
Db        660 TGGACACCTGAAACAGAAGTGGCTTTCCAGGCCC--TAAAGAAGTCTCTAACCCAAGACT 603

Qy       3409 CAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG-AAAAAAACAGR 3467
              ||||||||||:||||||||:|||||||||||||||:||:|||:: ||||| ||||||||||:
Db        602 CAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATAACACAGAAAAAAAACAGG 543

Qy       3468 AAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCT 3527
              ||:|||||:|||||||||||||||:|||:|||||:||||||||||||::||||::||||
Db        542 AATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGATGAGCTTGCAACCCGTGGCATACCT 483

Qy       3528 GASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGC 3587
              ||:||||||||:|||||||||||||||||||||:|:|||| |||||: ||||||||  ||
Db        482 GAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATGGTTTACAGGTAGTGGCAGC 423

Qy       3588 AGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTG 3647
              |||||||| ::|||||||||||| |||||||| |||||| ||:|||||||||||||||||
Db        422 AGTAGCAGGCTTAGTATCTGAAGAAGTTAAAAGAATACAAGGAAGAGATCTTACTGTGTG 363
```

```
Qy       3648 GACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAA 3707
              |:|||| ||||:|||||::|||||||||||||||||||||||||||||||||||||||||
Db        362 GRCATC-CATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAA 304

Qy       3708 CYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTG 3767
              |:  ||||||||||:|||||||||||||||||||: |||||||||||:|||| |  ||||
Db        303 CCACTTACTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCAACTGCGTGCTTG 244

Qy       3768 TGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTG 3827
              |||||||||||||  ||:| |||||||||||| | ||| ||||||||||:||:|||||||
Db        243 TGCAACTCTTAACTGAGCCACATTTCTTCCAAATAATAAAGAAAAGATAGAACATAACTG 184

Qy       3828 TCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGAC 3887
              |||||||:|||||  ||||||||||  |||||||||||||  |||:||||:|||||:|||||
Db        183 TCAACAAGTAATTGCTCAAACCTACACCACTCGAGGGGAACTTTTAGAGGTTCCCTTGAC 124

Qy       3888 TGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAA 3947
              ||||||:|| ||||||||||||  |||||||||||||||  |||||||||||||||||| ||
Db        123 TGATCCTGA-CTTCAACTTGTACACTGATGGAAGTTCCGTTGTAGAAAAAGGACTTCAAA 65

Qy       3948 AAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAG 4007
              |||:| ||||  |||||||||||||||||||||:|||||||||||| |||||| ||||
Db         64 AAGCGAGGTACACAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCTCCTCACACCAG 5

Qy       4008 GAAC 4011
              ||||
Db          4 GAAC 1
```

RESULT 113
US-09-925-065A-937161/c
; Sequence 937161, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 937161
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-937161

Query Match 7.8%; Score 581.5; DB 1; Length 720;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 621; Conservative 51; Mismatches 47; Indels 5; Gaps 4;

```
Qy       3289 GCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGA 3348
              | ||| :|||||:|||||:|||||||||||||||||||| ||||| |||||||:|||||
Db        720 GACAGACCATTATATATACACTAATTAAGGAAACTCAGAAAGTCAATATCCATTTAGTAAGA 661

Qy       3349 TGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
              ||||:|:||||| :::|||||||||||||||||    ||  | | ||  |||||||| |
Db        660 TGGACACCTGAAACAGAAGTGGCTTTCCAGGCCC--TAAAGAAGTCTCTAACCCAAGACT 603

Qy       3409 CAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG-AAAAAAACAGR 3467
              ||||||||||:||||||||:|||||||||||||||:||:|||:: ||||| ||||||||||:
```

```
Db        602 CAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATAACACAGAAAAAAAACAGG 543

Qy       3468 AAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCT 3527
              ||:|||||||:|||||||||||||||:|||:|||||:||||||||||||::|||||::||||
Db        542 AATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGATGAGCTTGCAACCCGTGGCATACCT 483

Qy       3528 GASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGC 3587
              ||:||||||||:|||||||||||||||||||||:|:|||| |||||: |||||||| ||
Db        482 GAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATGGTTTACAGGTAGTGGCAGC 423

Qy       3588 AGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTG 3647
              |||||||| ::|||||||||||| |||||||| |||||| ||:||||||||||||||||||
Db        422 AGTAGCAGGCTTAGTATCTGAAGAAGTTAAAAGAATACAAGGAAGAGATCTTACTGTGTG 363

Qy       3648 GACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAA 3707
              |:|||| |||||:|||||::|||||||||||||||||||||||||||||||||||||||||
Db        362 GRCATC-CATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAA 304

Qy       3708 CYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTG 3767
              |: ||||||||||:|||||||||||||||||: ||||||||||:|||| |  ||||
Db        303 CCACTTACTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCAACTGCGTGCTTG 244

Qy       3768 TGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTG 3827
              |||||||||||||  ||:| |||||||||| | ||| ||||||||||:||:|||||||
Db        243 TGCAACTCTTAACTGAGCCACATTTCTTCCAAATAATAAAGAAAAGATAGAACATAACTG 184

Qy       3828 TCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGAC 3887
              |||||||:||||| |||||||||  |||||||||||| |||:||||:|||||:|||||
Db        183 TCAACAAGTAATTGCTCAAACCTACACCACTCGAGGGGAACTTTTAGAGGTTCCCTTGAC 124

Qy       3888 TGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAA 3947
              ||||||:|| |||||||||||| |||||||||||||||| ||||||||||||||||| ||
Db        123 TGATCCTGA-CTTCAACTTGTACACTGATGGAAGTTCCGTTGTAGAAAAAGGACTTCAAA 65

Qy       3948 AAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAG 4007
              |||:| ||||  |||||||||||||||||||||:|||||||||||| |||||| ||||
Db          64 AAGCGAGGTACACAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCTCCTCACACCAG 5

Qy       4008 GAAC 4011
              ||||
Db           4 GAAC 1


RESULT 114
US-09-925-065A-940856/c
; Sequence 940856, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 940856
;   LENGTH: 803
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-940856

  Query Match             7.7%;  Score 579.7;  DB 1;  Length 803;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  721;  Conservative   15;  Mismatches   65;  Indels   25;  Gaps   10;

Qy        801 AAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGT 860
              |||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        803 AAGTGCCATATGTACAAACTTTCTTTTCATTAAGAGACAACTCGCAATTATGTAAAAAGT 744

Qy        861 GTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCAT--CCCCGA 918
              |||||||||||||||||||||||| |||||||||||||||||| |||||| |  ||||||
Db        743 GTGATTTATGCCCTACAGGAAGCCCTCAGAGTCTACCTCCCTACCCCAGCGTCCCCCCGA 684

Qy        919 CTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACA 978
              ||||||||||||::|||||||||||||||| |||||||||| ||| ||||||||||||||||
Db        683 CTCCTTCCCCAACTAATAAGGACCCCCTTTCAACCCAAACGGTACAAAAGGAGATAGACA 624

Qy        979 AAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCA 1038
              || |||||||||  |||||||||||||||||||||||||||| |||||| || || |||||||
Db        623 AAGGGGTAAACAACAAACCAAAGAGTGCCAATATTCCCCGATTATGCCCACT-CCAAGCA 565

Qy       1039 GTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTA 1098
              |||||||| ||||||| || |||||||||||||||| |:||||::|||| |||||||
Db        564 GTGGGAGGAGGAGAATTTGGTCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAGACTT- 506

Qy       1099 AAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTT 1158
                    || |||| |||||||||||||| ||||:|||||||||||||| ||:|| |
Db        505 ----------AATAGACCTAGGTAAATTCTCATATAACCCTGATGGCTATACTGATGCT 457

Qy       1159 TTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAA 1218
              |||||||||||||||||| |||||||||| |||||||||||||| |||| |||||||||
Db        456 TTACAAGGGTTAGGACAATCCTTTGATCTGATATGGAGAGATATA-ATGTTACTGCTAAA 398

Qy       1219 TCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGA 1278
              |||||||||||:|||||||||||||||||||| ||||||||||||||  ||:||||||| |
Db        397 TCAGACACTAACYCCAAATGAGAGAAGTGCTGCCATAACTGCAGCCCAAGAGTTTGGCAA 338

Qy       1279 TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGATT 1338
              |||||||||| ||||| |||||||||      |||||||||||| |||||||| ||  |||
Db        337 TCTCTGGTATTTCAGTTAGGTCAATG--ATAGGATGACAACAG-AGGAAAGAGAACAATT 281

Qy       1339 CCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTA 1398
              ||||||||||||||:|||||||||||| ||:|||||||||||||   ||| | |  |
Db        280 CCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG---ACACAGAATCAGA 224

Qy       1399 ACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAA 1458
              |||| |||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        223 ACAT-GGAGATTGGTGCCACAGACATTTGCTAACTTGCGTGCTAGAAGGACTAAGGAAAA 165

Qy       1459 CTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGA 1518
              |||:|||| ||: |||:|||||| ||| |||||||||| |||||||| ||||| ||||||
Db        164 CTAGGAAG-AAGCCTATGAATTATTCACTGATGTCCACTATAACACA-GGGAAAGGAAGA 107

Qy       1519 AAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTC 1578
              |||||||| |||||||||||||||||||||||||||||||||| |||| | ||||||||||
Db        106 AAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGAAAGCATACCTCTCTGTC 47

Qy       1579 ACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATC 1624
              |||||||||| |||||||||||||||||||| ||||||||||
Db          46 ACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAGTTTATC 1


RESULT 115
US-09-925-065A-940856/c
; Sequence 940856, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 940856
;   LENGTH: 803
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-940856

  Query Match             7.7%;  Score 579.7;  DB 1;  Length 803;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  721;  Conservative   15;  Mismatches   65;  Indels   25;  Gaps   10;

Qy        801 AAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGT 860
              |||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||||
Db        803 AAGTGCCATATGTACAAACTTTCTTTTCATTAAGAGACAACTCGCAATTATGTAAAAAGT 744

Qy        861 GTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCAT--CCCCGA 918
              |||||||||||||||||||||||| |||||||||||||||||| |||||| |  ||||||
Db        743 GTGATTTATGCCCTACAGGAAGCCCTCAGAGTCTACCTCCCTACCCCAGCGTCCCCCCGA 684

Qy        919 CTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACA 978
              ||||||||||::|||||||||||||| |||||||||| ||| ||||||||||||||||
Db        683 CTCCTTCCCCAACTAATAAGGACCCCCTTTCAACCCAAACGGTACAAAAGGAGATAGACA 624

Qy        979 AAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCA 1038
              || |||||||||   |||||||||||||||||||||||| |||||| || || |||||||
Db        623 AAGGGGTAAACAACAAACCAAAGAGTGCCAATATTCCCCGATTATGCCCACT-CCAAGCA 565

Qy       1039 GTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTA 1098
              ||||||||| ||||||| || |||||||||||||||||| |:||||::|||| |||||||
Db        564 GTGGGAGGAGGAGAATTTGGTCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAGACTT- 506

Qy       1099 AAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTT 1158
                    || |||| |||||||||||||| ||||:|||||||||||||| ||:|| |
Db        505 -----------AATAGACCTAGGTAAATTCTCATATAACCCTGATGGCTATACTGATGCT 457

Qy       1159 TTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAA 1218
              |||||||||||||||||| |||||||||| |||||||||||||| |||| |||||||||
Db        456 TTACAAGGGTTAGGACAATCCTTTGATCTGATATGGAGAGATATA-ATGTTACTGCTAAA 398

Qy       1219 TCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGA 1278
              ||||||||||||:||||||||||||||||  ||||||||||||||  ||:||||||| |
Db        397 TCAGACACTAACYCCAAATGAGAGAAGTGCTGCCATAACTGCAGCCCAAGAGTTTGGCAA 338

Qy       1279 TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGATT 1338
              |||||||||| ||||| |||||||||     |||||||||||| |||||||| ||  |||
Db        337 TCTCTGGTATTTCAGTTAGGTCAATG--ATAGGATGACAACAG-AGGAAAGAGAACAATT 281

Qy       1339 CCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTA 1398
              ||||||||||||||:|||||||||||| ||:||||||||||||   ||| | |     |
Db        280 CCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG---ACACAGAATCAGA 224

Qy       1399 ACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAA 1458
              |||| |||||||||||   |||||||||||||||||| ||||||:|||||||||||||||
Db        223 ACAT-GGAGATTGGTGCCACAGACATTTGCTAACTTGCGTGCTAGAAGGACTAAGGAAAA 165

Qy       1459 CTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGA 1518
              |||:|||| ||: |||:|||||| ||| |||||||||| |||||||| ||||| ||||||
Db        164 CTAGGAAG-AAGCCTATGAATTATTCACTGATGTCCACTATAACACA-GGGAAAGGAAGA 107

Qy       1519 AAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTC 1578
              ||||||||| |||||||||||||||||||||||||||||||| |||| | ||||||||||
Db        106 AAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGAAAGCATACCTCTCTGTC 47

Qy       1579 ACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATC 1624
              |||||||||| ||||||||||||||||||||||  ||||||||||
Db         46 ACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAGTTTATC 1


RESULT 116
US-08-979-847-112
; Sequence 112, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  112:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-112

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
              ||||| |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
              ||||||||||:||    |||||| |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
              ||||| | ||||||||||:||||||:|||||||||||||| |: :||:|||||||||||
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
              |||||||||||| | ||||| |||||||||||| |||||||||||||| |:||||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236
```

```
Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
             ||||| |||||||  :|| | |  || |||||||||||||| ||||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
             ||||||    |||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| |||| |||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
             |||||||||| ||||| |||||||| |||||  | || ||| ||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
             |||||||   |||||||||||| |||||| |||||||| |   |||| ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
             |||||| |||||  |||||||   |||| | ||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
             |||||| |||||| |||| ||| |||||||||||||||||||||||  | | |  |||||
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
             ||||| |||| |||| |||||||| |||||||||||||: ||||||||
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 117
US-10-114-104-112
; Sequence 112, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 112:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 758 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-10-114-104-112

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||| ||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:||    |||||| |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
             ||||| | ||||||||||:||||||:|||||||||||||||| |: :||:||||||||||
Db        117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
             |||||||||||||| | ||||| |||||||||||||| |||||||||||||||| |:|||||
Db        177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
             ||||| |||||||  :|| | |  || |||||||||||||| ||||||||||||||||||
Db        237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
             ||||||    |||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db        297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
              |||| |||||| |||| |||||||||||||||||||||| |||||| |||| ||||
Db        353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
             |||||||||| ||||| |||||||| |||||  | || ||| ||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
             |||||||   |||||||||||| |||||| |||||||| |   |||| ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
             |||||| |||||  |||||||   |||| | ||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
             |||||| |||||| |||| ||| |||||||||||||||||||||||  | | |  |||||
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
             ||||| |||| |||| |||||||| |||||||||||||: ||||||||
```

```
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 118
US-09-925-065A-734747/c
; Sequence 734747, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 734747
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-734747

  Query Match             7.4%;  Score 553.3;  DB 1;  Length 632;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  583;  Conservative   45;  Mismatches    4;  Indels    9;  Gaps    5;

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
              :|||||||||||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||
Db        632 GTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGATACCTGAAGCAAAA 573

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
              |||||||||||||| |||||||||||||||| ||||||||:|||||||||||||:|||||||
Db        572 GTGGCTTTCCAGG-CCCTAAAGAAGGCCTT-AACCCAAGTCCCAGTGTTAAGCTTGCCAA 515

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
              |:|||||||||||||:||:|||::|||||||||||||||:||:||||||:|||||||||
Db        514 CGGGGCAAGACTTTTCTTTATACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTT 455

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
              ||||||:|||:|||||: |||||||||||::||||::||||||:|||||||:|||||||
Db        454 ACACAGGTCCAAGGGACAAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTA 395

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
              |||||||||||||:|:||||||||||:||||||||||||||||||||:|||||||||
Db        394 GTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGTGGCAGTAGCAGTTKTAGTATCT 335

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              |||||||||||||||||||| |||||||||||||||||||||||||:|||||:
Db        334 GAAGCAGTTAAAATAATACAGGG-AGAGATCTTACTGTGTGGACATCTCATGAGGTGAAC 276

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :||||||||||||||||||||||||||||||||||||||||:|||||||||||||:||||
Db        275 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAG 216

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |||||||||||||:||||||||||:||||||||||||||||||||||||||||||:|
Db        215 GCTCTATTACTTGAAAGGCCAGTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTC 156

Qy       3787 NCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAA 3846
              |||||||||||||||||     ||||||:||:||||||||||||:||||||||||
Db        155 ACATTTCTTCCAGACAATG-----AAGATAGAATATAACTGTCAACAAATAATTTCTCAA 101

Qy       3847 ACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTT 3906
              |||||||||||||||||||||||:||||:|||||:||||||||||:|||| |||||||
Db        100 ACCTATGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTT 42

Qy       3907 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAA 3947
              |||||||||||||||||||||||||||||||||||| || ||
Db         41 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACCTCAAA 1


RESULT 119
US-09-925-065A-734747/c
; Sequence 734747, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 734747
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-734747

  Query Match             7.4%;  Score 553.3;  DB 1;  Length 632;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  583;  Conservative   45;  Mismatches    4;  Indels    9;  Gaps    5;

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
              :|||||||||||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||
Db        632 GTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGATACCTGAAGCAAAA 573

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
              |||||||||||||| |||||||||||||||| ||||||||:|||||||||||||:|||||||
Db        572 GTGGCTTTCCAGG-CCCTAAAGAAGGCCTT-AACCCAAGTCCCAGTGTTAAGCTTGCCAA 515

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
              |:|||||||||||||:||:|||::|||||||||||||||:||:||||||:|||||||||
Db        514 CGGGGCAAGACTTTTCTTTATACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTT 455

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
              ||||||:|||:|||||: |||||||||||::||||::||||||:|||||||:|||||||
Db        454 ACACAGGTCCAAGGGACAAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTA 395

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
              |||||||||||||:|:||||||||||:||||||||||||||||||||:|||||||||
Db        394 GTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGTGGCAGTAGCAGTTKTAGTATCT 335

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              |||||||||||||||||||| |||||||||||||||||||||||||:|||||:
Db        334 GAAGCAGTTAAAATAATACAGGG-AGAGATCTTACTGTGTGGACATCTCATGAGGTGAAC 276

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :||||||||||||||||||||||||||||||||||||||||:|||||||||||||:||||
Db        275 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAG 216

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |||||||||||||:||||||||||:||||||||||||||||||||||||||||||:|
Db        215 GCTCTATTACTTGAAAGGCCAGTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTC 156
```

```
Qy       3787 NCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAA 3846
              ||||||||||||||||||     ||||||:||:|||||||||||||:|||||||||||
Db        155 ACATTTCTTCCAGACAATG-----AAGATAGAATATAACTGTCAACAAATAATTTCTCAA 101

Qy       3847 ACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTT 3906
              ||||||||||||||||||||||||:||||:|||||:||||||||||:|||| |||||||
Db        100 ACCTATGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTT 42

Qy       3907 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAA 3947
              |||||||||||||||||||||||||||||||||| || ||
Db         41 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACCTCAAA 1

RESULT 120
US-09-925-065A-956040
; Sequence 956040, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956040
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956040

  Query Match             7.4%;  Score 553.2;  DB 1;  Length 752;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  638;  Conservative    3;  Mismatches   72;  Indels    8;  Gaps    4;

Qy       1652 AACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAA 1711
              |||||||||||||||| ||||||| |||||||||||||||||||| ||||||||||||
Db          1 AACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAA 60

Qy       1712 CYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAA 1771
              |:|| |||||||||||||||||||||||||||||| ||||   ||||||   |  | |:
Db         61 CTTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAM 120

Qy       1772 AAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAA 1831
              |||||||||||| |||||||||||| ||||||||||||||||||||||||||||| ||||||
Db        121 AAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAA 180

Qy       1832 GGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 1891
              ||||||  |||| ||||| ||||||| ||||||||||||||||||||||||||||||||
Db        181 GGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 240

Qy       1892 CACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTT 1951
              |||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||| |
Db        241 CACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGT 300

Qy       1952 CAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCC 2011
              |||||||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||
Db        301 CAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCC 360

Qy       2012 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCC 2071
              |||||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||
Db        361 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCC 420

Qy       2072 ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTC 2128
              ||||||||||||||| |||  |||||||||||||||||||||||||||| | |   ||||||
Db        421 ATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTC 480

Qy       2129 CTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGA 2188
              |||||||||||||| | ||||| |||||||||| ||||||| ||||||||||||||||||
Db        481 CTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGA 540

Qy       2189 TCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCC 2248
              ||||||| ||| |     |||  | || ||| ||||||||||||||  ||:||||||||
Db        541 TCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCC 596

Qy       2249 ACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAA 2308
              |||||||| | || |||||| ||||| |||||||||||| |||||||||||  |||||
Db        597 ACTAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAA 655

Qy       2309 GCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGA 2368
              |||||| | | || |||||||| ||||| |||||||||||||||||||||||||||  ||
Db        656 GCCCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGA 715

Qy       2369 A 2369
              |
Db        716 A 716

RESULT 121
US-09-925-065A-956040
; Sequence 956040, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956040
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956040

  Query Match             7.4%;  Score 553.2;  DB 1;  Length 752;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  638;  Conservative    3;  Mismatches   72;  Indels    8;  Gaps    4;

Qy       1652 AACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAA 1711
              |||||||||||||||| ||||||| |||||||||||||||||||| ||||||||||||
Db          1 AACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAA 60

Qy       1712 CYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAA 1771
              |:|| |||||||||||||||||||||||||||||| ||||   ||||||   |  | |:
Db         61 CTTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAM 120

Qy       1772 AAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAA 1831
              |||||||||||| |||||||||||| ||||||||||||||||||||||||||||| ||||||
Db        121 AAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAA 180

Qy       1832 GGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 1891
              ||||||  |||| ||||| ||||||| ||||||||||||||||||||||||||||||||
```

```
Db        181 GGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 240

Qy       1892 CACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTT 1951
              |||||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||| |
Db        241 CACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGT 300

Qy       1952 CAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCC 2011
              |||||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||
Db        301 CAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCC 360

Qy       2012 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCC 2071
              |||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||
Db        361 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCC 420

Qy       2072 ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTC 2128
              ||||||||||||| |||  |||||||||||||||||||||||||| | |   ||||||
Db        421 ATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTC 480

Qy       2129 CTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGA 2188
              |||||||||||||| | |||||  |||||||||| |||||||  |||||||||||||||||
Db        481 CTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGA 540

Qy       2189 TCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCC 2248
              ||||||| ||| |     |||  | || ||| ||||||||||||||  ||:||||||||
Db        541 TCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCC 596

Qy       2249 ACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAA 2308
              |||||||| | || |||||| ||||| |||||||||||| ||||||||||||  |||||
Db        597 ACTAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAA 655

Qy       2309 GCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGA 2368
              |||||| | | || |||||||| ||||| |||||||||||||||||||||||||||  ||
Db        656 GCCCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGA 715

Qy       2369 A 2369
              |
Db        716 A 716
```

RESULT 122
US-09-925-065A-956039
; Sequence 956039, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956039
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-956039

Query Match 7.4%; Score 551.6; DB 1; Length 752;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 637; Conservative 3; Mismatches 73; Indels 8; Gaps 4;

```
Qy       1652 AACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAA 1711
              |||||||||||||||| ||||||| |||||||||||||||||||| ||||||||||||
Db          1 AACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAA 60

Qy       1712 CYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAA 1771
              |:|| ||||||||:|||||||||||||||||||||| ||||   ||||||   |  | |
Db         61 CTTCAGTTTTTTAYAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAC 120

Qy       1772 AAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAA 1831
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||| ||||||
Db         121 AAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAA 180

Qy       1832 GGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 1891
              |||||| |||| ||||| ||||||| ||||||||||||||||||||||||||||||||||
Db         181 GGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 240

Qy       1892 CACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTT 1951
              |||||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||| |
Db         241 CACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGT 300

Qy       1952 CAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCC 2011
              |||||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||
Db         301 CAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCC 360

Qy       2012 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCC 2071
              |||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||
Db         361 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCC 420

Qy       2072 ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTC 2128
              ||||||||||||| |||  |||||||||||||||||||||||||| | |   ||||||
Db         421 ATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTC 480

Qy       2129 CTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGA 2188
              |||||||||||||| | |||||  |||||||||| |||||||  |||||||||||||||||
Db         481 CTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGA 540

Qy       2189 TCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCC 2248
              ||||||| ||| |     |||  | || ||| ||||||||||||||  ||:||||||||
Db         541 TCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCC 596

Qy       2249 ACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAA 2308
              |||||||| | || |||||| ||||| |||||||||||| ||||||||||||  |||||
Db         597 ACTAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAA 655

Qy       2309 GCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGA 2368
              |||||| | | || |||||||| ||||| |||||||||||||||||||||||||||  ||
Db         656 GCCCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGA 715

Qy       2369 A 2369
              |
Db         716 A 716
```

RESULT 123
US-09-925-065A-956039
; Sequence 956039, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

```
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956039
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956039

  Query Match             7.4%;  Score 551.6;  DB 1;  Length 752;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  637;  Conservative    3;  Mismatches   73;  Indels    8;  Gaps    4;

Qy       1652 AACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAA 1711
             |||||||||||||||| ||||||| |||||||||||||||||| ||||||||||||
Db             1 AACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAA 60

Qy       1712 CYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAA 1771
             |:|| |||||||:|||||||||||||||||||| ||||    ||||||   |  | |
Db            61 CTTCAGTTTTTTAYAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAC 120

Qy       1772 AAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAA 1831
             |||||||||||| |||||||||| |||||||||||||||||||||||||||| ||||||
Db        121 AAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAA 180

Qy       1832 GGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 1891
             ||||||  |||| |||||  ||||||| ||||||||||||||||||||||||||||||||||
Db        181 GGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGA 240

Qy       1892 CACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTT 1951
             |||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||| |
Db        241 CACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGT 300

Qy       1952 CAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCC 2011
             |||||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||
Db        301 CAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCC 360

Qy       2012 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCC 2071
             |||||||||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||
Db        361 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCC 420

Qy       2072 ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTC 2128
             |||||||||||||| |||  |||||||||||||||||||||||||| | |   ||||||
Db        421 ATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTC 480

Qy       2129 CTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGA 2188
             |||||||||||||| | ||||| |||||||||| ||||||| ||||||||||||||||||
Db        481 CTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGA 540

Qy       2189 TCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCC 2248
             ||||||| ||| |      |||  | || ||| ||||||||||||||  ||:||||||||
Db        541 TCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCC 596

Qy       2249 ACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAA 2308
             |||||||| | || |||||  |||||  |||||||||||||  |||||||||||  |||||
Db        597 ACTAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAA 655

Qy       2309 GCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGA 2368
             |||||| | | || |||||||| ||||| ||||||||||||||||||||||||||  ||
Db        656 GCCCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGA 715

Qy       2369 A 2369
             |
Db        716 A 716


RESULT 124
US-09-925-065A-948058/c
; Sequence 948058, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948058
;   LENGTH: 751
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948058

  Query Match             7.4%;  Score 551.2;  DB 1;  Length 751;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  636;  Conservative    3;  Mismatches   72;  Indels    8;  Gaps    4;

Qy       1654 CTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACY 1713
             |||||||||||||| ||||||| |||||||||||||||||| ||||||||||||:
Db         751 CTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAACT 692

Qy       1714 TCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAA 1773
             || ||||||||:|||||||||||||||||||| ||||   ||||||   |  | ||||
Db         691 TCAGTTTTTTAYAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAAAA 632

Qy       1774 AAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGG 1833
             |||||||||| |||||||||| |||||||||||||||||||||||||||| ||||||||
Db         631 AAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAAGG 572

Qy       1834 GAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACA 1893
             ||||  |||| ||||| ||||||| ||||||||||||||||||||||||||||||||||
Db         571 GAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACA 512

Qy       1894 CTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCA 1953
             |||||||||||||||||| | ||||| |||||| |||| |||||||||||||||||| |||
Db         511 CTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGTCA 452

Qy       1954 AGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCAC 2013
             |||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||||
Db         451 AGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCCAC 392

Qy       2014 TAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCAT 2073
             |||||||||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||||
Db         391 TAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCAT 332

Qy       2074 CACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCT 2130
             |||||||||||  |||  |||||||||||||||||||||||||||| | |   ||||||||
Db         331 CACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCT 272

Qy       2131 GGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATC 2190
             |||||||||||  | ||||| |||||||||| ||||||| ||||||||||||||||||
Db         271 GGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGATC 212

Qy       2191 TGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCAC 2250
             ||||| ||| |      |||  | || ||| ||||||||||||||  ||:||||||||||
Db         211 TGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCCAC 156

Qy       2251 TAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGC 2310
             |||||| | || |||||| ||||| |||||||||||||  |||||||||||  |||||||
Db         155 TAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAAGC 97

Qy       2311 CCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAA 2369
             |||| | | || |||||||| ||||| ||||||||||||||||||||||||||  |||
Db          96 CCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGAA 38
```

```
RESULT 125
US-09-925-065A-948058/c
; Sequence 948058, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948058
;   LENGTH: 751
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948058

  Query Match             7.4%;  Score 551.2;  DB 1;  Length 751;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  636;  Conservative    3;  Mismatches   72;  Indels    8;  Gaps    4;

Qy       1654 CTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACY 1713
             |||||||||||||||| ||||||| |||||||||||||||||||| ||||||||||||||:
Db        751 CTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTGGCAACT 692

Qy       1714 TCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAA 1773
             || ||||||||:|||||||||||||||||||||| ||||   ||||||   |  | ||||
Db        691 TCAGTTTTTTAYAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGATTAAAA 632

Qy       1774 AAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGG 1833
             |||||||||| |||||||||| |||||||||||||||||||||||||||||| ||||||||
Db        631 AAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGGAAAAGG 572

Qy       1834 GAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACA 1893
             ||||  |||| ||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db        571 GAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACA 512

Qy       1894 CTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCA 1953
             |||||||||||||||| | ||||| |||||| |||| ||||||||||||||| |||
Db        511 CTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGTCA 452

Qy       1954 AGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCAC 2013
             |||||||||||||||||||||||||||| |||||||  |||||  | ||||||||||||||||
Db        451 AGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGAAGCCAC 392

Qy       2014 TAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCAT 2073
             |||||||||||||||||||||||||||||||||||| |||||||| |||| ||||||||||
Db        391 TAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCAT 332

Qy       2074 CACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCT 2130
             |||||||||||| |||  |||||||||||||||||||||||||| | |   ||||||||
Db        331 CACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCT 272

Qy       2131 GGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATC 2190
             |||||||||||| | ||||| |||||||||| ||||||| ||||||||||||||||||||
Db        271 GGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTCCAGATC 212

Qy       2191 TGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCAC 2250
             ||||| ||| |     |||  | || ||| ||||||||||||||  ||:||||||||||
Db        211 TGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCCAGCCAC 156

Qy       2251 TAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGC 2310
             |||||| | || |||||| ||||| |||||||||||||  |||||||||||||  |||||||
Db        155 TAAGTTGT-AAATGGGGAACTTTACTCTTTTCACATGCCTTTCTAATTATGACTGAAAGC 97

Qy       2311 CCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAA 2369
             |||| | | || |||||||| ||||| |||||||||||||||||||||||||||  |||
Db         96 CCCATTCCTTTATTAGGGAGTGACATCCTAGCAAAAGCAGGGGCCATTATACACTAGAA 38


RESULT 126
US-09-925-065A-799227/c
; Sequence 799227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 799227
;   LENGTH: 626
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-799227

  Query Match             7.4%;  Score 550.5;  DB 1;  Length 626;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  579;  Conservative   45;  Mismatches    2;  Indels    9;  Gaps    5;

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
             :|||||||||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||
Db        626 GTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGATACCTGAAGCAAAA 567

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
             |||||||||||||| |||||||||||||||| ||||||||:||||||||||:|||||||
Db        566 GTGGCTTTCCAGG-CCCTAAAGAAGGCCTT-AACCCAAGTCCCAGTGTTAAGCTTGCCAA 509

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
             |:|||||||||||||:||:|||::|||||||||||||||:||:||||||:|||||||||
Db        508 CGGGGCAAGACTTTTCTTTATACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTT 449

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
             ||||||:|||:|||||: ||||||||||::||||::||||||:||||||||:|||||||
Db        448 ACACAGGTCCAAGGGACAAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTA 389

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
             |||||||||||||:|:||||||||||:|||||||||||||||||||||:|||||||||
Db        388 GTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGTGGCAGTAGCAGTTKTAGTATCT 329

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||:||||:
Db        328 GAAGCAGTTAAAATAATACAGGG-AGAGATCTTACTGTGTGGACATCTCATGAGGTGAAC 270

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
             :|||||||||||||||||||||||||||||||||||||||||||:|||||||||||:||||
Db        269 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAG 210

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             ||||||||||||||:|||||||||||:|||||||||||||||||||||||||||||||:|
```

```
Db        209 GCTCTATTACTTGAAAGGCCAGTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTC 150

Qy       3787 NCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAA 3846
              ||||||||||||||||||      ||||||:||:|||||||||||||:||||||||||
Db        149 ACATTTCTTCCAGACAATG-----AAGATAGAATATAACTGTCAACAAATAATTTCTCAA 95

Qy       3847 ACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTT 3906
              |||||||||||||||||||||||:||||:|||||:||||||||||:|||| |||||||
Db         94 ACCTATGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTT 36

Qy       3907 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
              |||||||||||||||||||||||||||||||||||
Db          35 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 1


RESULT 127
US-09-925-065A-799227/c
; Sequence 799227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 799227
;   LENGTH: 626
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-799227

  Query Match             7.4%;  Score 550.5;  DB 1;  Length 626;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  579;  Conservative   45;  Mismatches    2;  Indels    9;  Gaps    5;

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
              :|||||||||||||||||||||||||||||||||||||:|||||||||:|:|||||||:||
Db        626 GTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGATACCTGAAGCAAAA 567

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
              |||||||||||||| |||||||||||||||| ||||||||:||||||||||||:|||||||
Db        566 GTGGCTTTCCAGG-CCCTAAAGAAGGCCTT-AACCCAAGTCCCAGTGTTAAGCTTGCCAA 509

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
              |:|||||||||||||:||:|||::|||||||||||||||:||:||||||:||||||||
Db        508 CGGGGCAAGACTTTTCTTTATACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTT 449

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
              ||||||:|||:|||||: |||||||||||::||||::||||||:||||||||:|||||||
Db        448 ACACAGGTCCAAGGGACAAGCTTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTA 389

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
              |||||||||||||:|:||||||||||:||||||||||||||||||||:|||||||||||
Db        388 GTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGTGGCAGTAGCAGTTKTAGTATCT 329

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              |||||||||||||||||||| |||||||||||||||||||||||||||||:|||||:
Db        328 GAAGCAGTTAAAATAATACAGGG-AGAGATCTTACTGTGTGGACATCTCATGAGGTGAAC 270

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :|||||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||
Db        269 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAG 210

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |||||||||||||||:|||||||||||:|||||||||||||||||||||||||||||:|
Db        209 GCTCTATTACTTGAAAGGCCAGTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTC 150

Qy       3787 NCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAA 3846
              ||||||||||||||||||      ||||||:||:|||||||||||||:||||||||||
Db        149 ACATTTCTTCCAGACAATG-----AAGATAGAATATAACTGTCAACAAATAATTTCTCAA 95

Qy       3847 ACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTT 3906
              |||||||||||||||||||||||:||||:|||||:||||||||||:|||| |||||||
Db         94 ACCTATGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTT 36

Qy       3907 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
              |||||||||||||||||||||||||||||||||||
Db          35 GTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 1


RESULT 128
US-10-301-480-299445/c
; Sequence 299445, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 299445
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-299445

  Query Match             7.1%;  Score 535.5;  DB 1;  Length 663;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  566;  Conservative   58;  Mismatches   38;  Indels    5;  Gaps    4;

Qy       2844 TTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAA 2903
              |||||||||||||||| |||| ||||||:  |||||||||||||||||:||||||||||
Db        662 TTTGCCTTTGAAGATCCTTCGAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAG 603

Qy       2904 GGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYM 2963
              |||||||| |||||:|||||||||||||||||||||||||||||||||||||:||:|::
Db        602 GGGTTCAGAGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTC 543

Qy       2964 TCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCA 3023
              |||||||||||||||||||||||| :|||  |||||||||||||||:||: ||::||||
Db        542 TCATACCTGGACACTCTTGTCCTTTGGTACATGGATGATTTACTTTTAGCCACCCATTCA 483

Qy       3024 GAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTAC 3083
              ||| |||||||||||||||||||||||:|||||:||||||||||:||||||||||||
Db        482 GAAGCCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTAC 423

Qy       3084 AWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARA 3143
              |: | |||||||:|:|:|||||:|||||||||||| || |||||||||||||:||||:|
Db        422 AA-GATTCCAAACCAAAGGCTCAGCTCTGCTCACAGAAGTTTAAATACTTAGGGCTAAAA 364

Qy       3144 TTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCT 3203
              |||||:||||| ||||:|||||||||||| ||||:::|| |||||||||||||| ||||
Db        363 TTATCSAAAGGTACCAGGGCCCTCAGTGGGGAACATATGCAGCCTATACTGGCTTGTCCT 304

Qy       3204 CATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWA 3263
              ||||:|||||||||| ||||||||||:|::|||||||||:|||:|||:|||||||  ||:|
```

```
Db        303 CATCCCAAAACCCT-AAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCAAATA 245

Qy       3264 TGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCA 3323
              |||||| ||||||||:||||:|||||||||| :|||||:|||||:|||||||||||||||
Db        244 TGGATT-CCCAGGTATGGCAAAATAGCCAGATCATTATATACACTAATTAAGGAAACTCA 186

Qy       3324 GAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCC 3383
              |||||||||||||||||:|||| ||||:|:|||||:: ||| ||||||||||| |
Db        185 GAAAGCCAATACCCATTTAGTAAGGTGGACACCTGAAGCACAAGCAGCTTTCCAGGCTC- 127

Qy       3384 TAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTST 3443
                ||  | | ||  ||||||||: |||||||| |:|||||| :|||||||| ||||:|
Db        126 -TAAAGAAGGCTCTAACCCAAGTTCCAGTGTTACGCTTGCCAAAGGGGCAAGATTTTTCT 68

Qy       3444 TYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAY 3503
              |:| |::|||||||||||||||:||:||||||:|||||||||||||||:|||:|||||:
Db         67 TTAAATGTCACAGAAAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGAT 8

Qy       3504 GAGCTTG 3510
              || ||||
Db           7 GAACTTG 1


RESULT 129
US-10-301-480-912854/c
; Sequence 912854, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 912854
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-912854

  Query Match             7.1%;  Score 535.5;  DB 1;  Length 663;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  566;  Conservative   58;  Mismatches   38;  Indels    5;  Gaps    4;

Qy       2844 TTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAA 2903
              |||||||||||||| |||| ||||||:  |||||||||||||||||:||||||||||
Db        662 TTTGCCTTTGAAGATCCTTCGAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAG 603

Qy       2904 GGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYM 2963
              |||||||| |||||:|||||||||||||||||||||||||||||||||||||:||:|::
Db        602 GGGTTCAGAGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTC 543

Qy       2964 TCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCA 3023
              |||||||||||||||||||||||| :|||  |||||||||||||||:||: ||::||||
Db        542 TCATACCTGGACACTCTTGTCCTTTGGTACATGGATGATTTACTTTTAGCCACCCATTCA 483

Qy       3024 GAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTAC 3083
              ||| |||||||||||||||||||||||:|||||:||||||||||:||||||||||||
Db        482 GAAGCCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTAC 423

Qy       3084 AWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARA 3143
              |: | ||||||||:|:|:|||||:|||||||||||| || ||||||||||||||:||||:|
Db        422 AA-GATTCCAAACCAAAGGCTCAGCTCTGCTCACAGAAGTTTAAATACTTAGGGCTAAAA 364

Qy       3144 TTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCT 3203
              |||||:|||||| ||||:|||||||||| ||||:::|| |||||||||||||||| ||||
Db        363 TTATCSAAAGGTACCAGGGCCCTCAGTGGGGAACATATGCAGCCTATACTGGCTTGTCCT 304

Qy       3204 CATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWA 3263
              ||||:|||||||| |||||||||||:|::|||||||||:|||:|||:|||||||  ||:|
Db        303 CATCCCAAAACCCT-AAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCAAATA 245

Qy       3264 TGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCA 3323
              |||||| ||||||||:||||:|||||||||| :|||||:|||||:|||||||||||||||
Db        244 TGGATT-CCCAGGTATGGCAAAATAGCCAGATCATTATATACACTAATTAAGGAAACTCA 186

Qy       3324 GAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCC 3383
              |||||||||||||||||:|||| ||||:|:|||||:: ||| ||||||||||| |
Db        185 GAAAGCCAATACCCATTTAGTAAGGTGGACACCTGAAGCACAAGCAGCTTTCCAGGCTC- 127

Qy       3384 TAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTST 3443
                ||  | | ||  ||||||||: |||||||| |:|||||| :|||||||| ||||:|
Db        126 -TAAAGAAGGCTCTAACCCAAGTTCCAGTGTTACGCTTGCCAAAGGGGCAAGATTTTTCT 68

Qy       3444 TYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAY 3503
              |:| |::|||||||||||||||:||:||||||:|||||||||||||||:|||:|||||:
Db         67 TTAAATGTCACAGAAAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGAT 8

Qy       3504 GAGCTTG 3510
              || ||||
Db           7 GAACTTG 1


RESULT 130
US-09-925-065A-359698/c
; Sequence 359698, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359698
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-359698

  Query Match             7.1%;  Score 532.7;  DB 1;  Length 636;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  549;  Conservative   34;  Mismatches   44;  Indels    3;  Gaps    1;

Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
              |||||:|||:|:|||||||||||||| |||||||:||||||||:||||||| |||||| ||
Db        633 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 574

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
              ||:|:|||||:|||||||| ||||||||||:: ||||||||:|:|||||||||||||||
Db        573 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 514

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
              |||||||:|||:||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        513 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 454

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
              |||||||||| :|||||||:|||||||||||||:|||||||||  |||||||||:|||
```

```
Db        453 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 394

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
             |||||||:|||:|||:|||||||| :||||||||||:|||||||  ||||||||||||
Db        393 GGGAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 334

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
             ||||||||||||||||||  ||    | | ||||||  ||||||||||  || ||||:||||
Db        333 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGARACCA 274

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
             | |||||||||||||||| ||||| |||||||||||||||| |||     ||||| ||||
Db        273 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 214

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
             ||||  || || ||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        213 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 154

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
             |||||||||||||||||| |||||||||||||||||| ||||||||||||||:||||
Db        153 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 94

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db         93 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 34

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||:||||||||:|:|||||||||  |||
Db         33 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 4
```

RESULT 131
US-09-925-065A-359698/c
; Sequence 359698, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359698
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-359698

Query Match 7.1%; Score 532.7; DB 1; Length 636;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
             |||||:|||:|:||||||||||||| |||||||:|||||||:|||||| ||
Db        633 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 574

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
             ||:|:||||||:|||||||| ||||||||||||:: |||||||:|:|||||||||||||
Db        573 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 514

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
             |||||||:|||:||||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        513 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 454

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
             |||||||||||| :|||||||:|||||||||||||:|||||||  ||||||||:|||
Db        453 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 394

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
             ||||||||:|||:|||:|||||||| :||||||||||:|||||||  ||||||||||||
Db        393 GGGAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 334

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
             ||||||||||||||||||  ||    | | ||||||  ||||||||||  || ||||:||||
Db        333 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGARACCA 274

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
             | |||||||||||||||| ||||| |||||||||||||||| |||     ||||| ||||
Db        273 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 214

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
             ||||  || || ||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        213 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 154

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
             |||||||||||||||||| |||||||||||||||||| ||||||||||||||:||||
Db        153 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 94

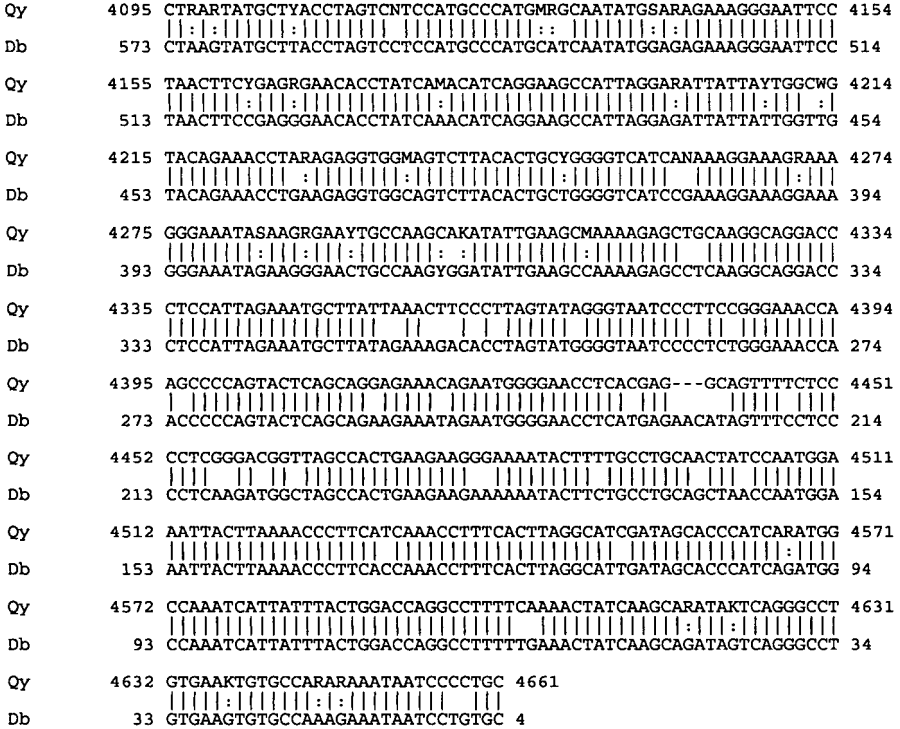
Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db         93 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 34

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||:||||||||:|:|||||||||  |||
Db         33 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 4
```

RESULT 132
US-09-925-065A-359699/c
; Sequence 359699, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359699
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-359699

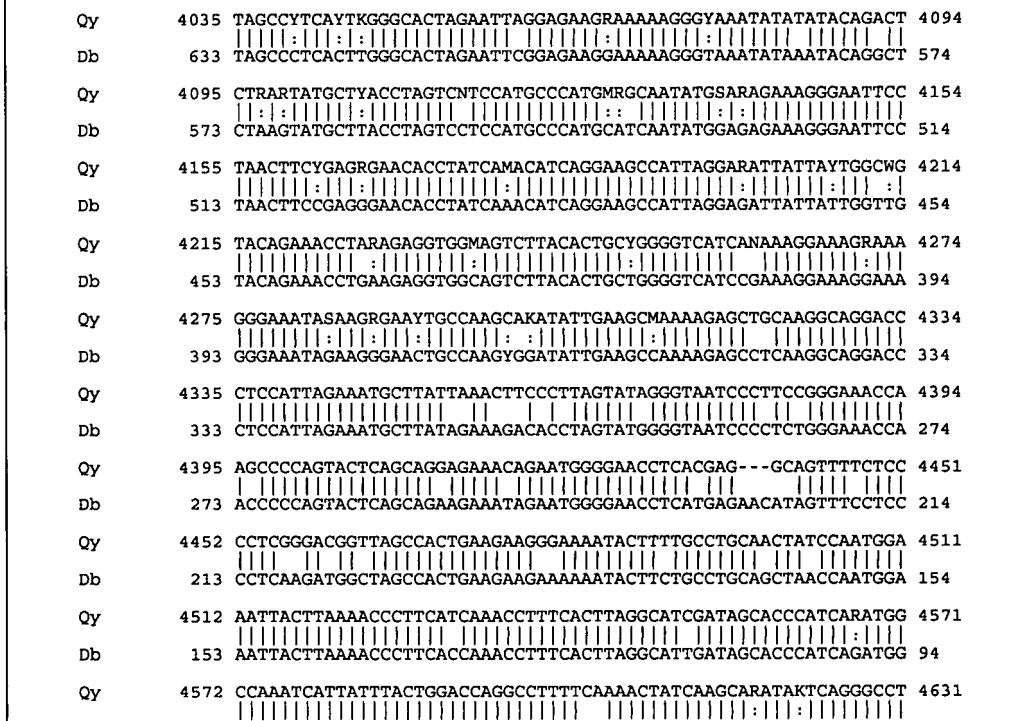
Query Match 7.1%; Score 532.7; DB 1; Length 636;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
             |||||:|||:|:||||||||||||| |||||||:|||||||:|||||| ||
Db        633 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 574
```



RESULT 133
US-09-925-065A-359699/c
; Sequence 359699, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359699
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-359699

Query Match 7.1%; Score 532.7; DB 1; Length 636;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;



RESULT 134
US-10-301-480-430862/c
; Sequence 430862, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 430862
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-430862

Query Match 7.1%; Score 532.7; DB 1; Length 637;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
             |||||:|||:|:|||||||||||||| |||||||:||||||||:||||||| |||||| ||
Db        634 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 575

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
             ||:|:|||||||:|||||||| |||||||||||:: |||||||:|:|||||||||||||| 
Db        574 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 515

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
             |||||||:|||:||||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        514 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 455

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
             |||||||||||| :|||||||:|||||||||||||:|||||||| ||||||||||:|||
Db        454 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 395

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
             ||||||||:|||:|||:|||||||| :|||||||||:|||||||  ||||||||||||
Db        394 GGGAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 335

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
             |||||||||||||||||  ||   | | |||||| |||||||||| || ||||:||||
Db        334 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGARACCA 275

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
             | ||||||||||||||||| ||||| |||||||||||||||| |||     |||||| ||||
Db        274 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 215

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
             ||||  || || ||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        214 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 155

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
             |||||||||||||||||| |||||||||||||||||| |||||||||||||:||||
Db        154 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 95

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db         94 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 35

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||:|||||||:|:|||||||||  |||
Db         34 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 5
```

RESULT 135
US-10-301-480-430863/c
; Sequence 430863, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 430863
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-430863

Query Match 7.1%; Score 532.7; DB 1; Length 637;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
             |||||:|||:|:|||||||||||||| |||||||:||||||||:||||||| |||||| ||
Db        634 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 575

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
             ||:|:|||||||:|||||||| |||||||||||:: |||||||:|:|||||||||||||| 
Db        574 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 515

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
             |||||||:|||:||||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        514 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 455

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
             |||||||||||| :|||||||:|||||||||||||:|||||||| ||||||||||:|||
Db        454 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 395

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
             ||||||||:|||:|||:|||||||| :|||||||||:|||||||  ||||||||||||
Db        394 GGGAAATAGAAGGGAACTGCCAAGYGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 335

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
             |||||||||||||||||  ||   | | |||||| |||||||||| || ||||:||||
Db        334 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGAAACCA 275

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
             | ||||||||||||||||| ||||| |||||||||||||||| |||     |||||| ||||
Db        274 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 215

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
             ||||  || || ||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        214 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 155

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
             |||||||||||||||||| |||||||||||||||||| |||||||||||||:||||
Db        154 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 95

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db         94 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 35

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||:|||||||:|:|||||||||  |||
Db         34 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 5
```

RESULT 136
US-10-301-480-1044271/c
; Sequence 1044271, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1044271
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1044271

Query Match 7.1%; Score 532.7; DB 1; Length 637;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
             |||||:|||:|:|||||||||||||| |||||||:||||||||:||||||| |||||| ||
```

```
Db        634 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 575

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
              ||:|:||||||:|||||||| ||||||||||:: |||||||:|:|||||||||||||||
Db        574 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 515

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
              |||||||:|||:||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        514 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 455

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
              |||||||||| :|||||||:||||||||||||:|||||||| ||||||||:|||
Db        454 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 395

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
              ||||||||:|||:|||:|||||||| :||||||||||:|||||||| ||||||||||||
Db        394 GGGAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 335

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
              |||||||||||||||||| || | | |||||| |||||||||| || ||||:||||
Db        334 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGARACCA 275

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
              | |||||||||||||||| ||||| |||||||||||||||| ||| |||||| ||||
Db        274 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 215

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
              ||||  || || |||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        214 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 155

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
              |||||||||||||||||| ||||||||||||||||||| ||||||||||||:||||
Db        154 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 95

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
              |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db          94 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 35

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
              |||||:|||||||:|:|||||||||  |||
Db          34 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 5
```

RESULT 137
US-10-301-480-1044272/c
; Sequence 1044272, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1044272
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1044272

Query Match 7.1%; Score 532.7; DB 1; Length 637;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 549; Conservative 34; Mismatches 44; Indels 3; Gaps 1;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACT 4094
              |||||:|||:|:|||||||||||||| |||||||:||||||||:|||||| |||||| ||
Db        634 TAGCCCTCACTTGGGCACTAGAATTCGGAGAAGGAAAAAGGGTAAATATAAATACAGGCT 575

Qy       4095 CTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCC 4154
              ||:|:||||||:|||||||| ||||||||||:: |||||||:|:|||||||||||||||
Db        574 CTAAGTATGCTTACCTAGTCCTCCATGCCCATGCATCAATATGGAGAGAAAGGGAATTCC 515

Qy       4155 TAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWG 4214
              |||||||:|||:|||||||||||:|||||||||||||||||||:|||||||:||| :|
Db        514 TAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGTTG 455

Qy       4215 TACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAA 4274
              |||||||||| :|||||||:||||||||||||:|||||||| ||||||||:|||
Db        454 TACAGAAACCTGAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCCGAAAGGAAAGGAAA 395

Qy       4275 GGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACC 4334
              ||||||||:|||:|||:|||||||| :||||||||||:|||||||| ||||||||||||
Db        394 GGGAAATAGAAGGGAACTGCCAAGYGGATATTGAAGCCAAAAGAGCCTCAAGGCAGGACC 335

Qy       4335 CTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCA 4394
              |||||||||||||||||| || | | |||||| |||||||||| || ||||||||||
Db        334 CTCCATTAGAAATGCTTATAGAAAGACACCTAGTATGGGGTAATCCCCTCTGGGAAACCA 275

Qy       4395 AGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTTTCTCC 4451
              | |||||||||||||||| ||||| |||||||||||||||| ||| |||||| ||||
Db        274 ACCCCCAGTACTCAGCAGAAGAAATAGAATGGGGAACCTCATGAGAACATAGTTTCCTCC 215

Qy       4452 CCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGA 4511
              ||||  || || |||||||||||||||  ||||||||| |||||||| ||| ||||||||
Db        214 CCTCAAGATGGCTAGCCACTGAAGAAGAAAAAATACTTCTGCCTGCAGCTAACCAATGGA 155

Qy       4512 AATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGG 4571
              |||||||||||||||||| ||||||||||||||||||| ||||||||||||:||||
Db        154 AATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGATGG 95

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
              |||||||||||||||||||||||||||||||  |||||||||||||:|||:|||||||||
Db          94 CCAAATCATTATTTACTGGACCAGGCCTTTTTGAAACTATCAAGCAGATAGTCAGGGCCT 35

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGC 4661
              |||||:|||||||:|:|||||||||  |||
Db          34 GTGAAGTGTGCCAAAGAAATAATCCTGTGC 5
```

RESULT 138
US-09-925-065A-926128/c
; Sequence 926128, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926128
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926128

Query Match 7.1%; Score 532.6; DB 1; Length 735;
Best Local Similarity 88.6%; Pred. No. 0;

```
  Matches  660;  Conservative    1;  Mismatches   66;  Indels   18;  Gaps    8;

Qy       1958 AATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAAC 2017
             ||||||||||||||||||||||||||||||||  ||||||  | ||||||||||||||||||
Db        735 AATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTAAC 676

Qy       2018 CAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACC 2077
             |||||  ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        675 CAGATTTTCCCGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACC 616

Qy       2078 CTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGAC 2134
             |||||||||||| ||||| |||||||||||||||||||||||| | |   ||||||||||||
Db        615 CTCACAGAGCCCCGGGTACGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGGAC 556

Qy       2135 ACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTC 2194
             ||||| |  || |||||  |||||||||||  ||||||||||||||||||||||| |||||||||
Db        555 ACTGGAGTGGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCCTCTAGATCTGTC 496

Qy       2195 ACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAG 2254
             |||||  |   | | |  | || ||| | ||||||||||||||| ||:|||||||||| |
Db        495 ACTATCC---GAGCGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTGAC 440

Qy       2255 TTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCA 2314
             || || ||||| || ||||| ||||| |||||||||||||||||||| ||||||||||
Db        439 TTGTG-ACTGGAGAACTTTACTCTTTCCACATGCTTTTCTAATTATGCCTGAAAGCCCCA 381

Qy       2315 CTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAG 2374
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CTCCCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAG 321

Qy       2375 GAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGG 2434
             ||||||||||||| |||||||| |||||||||||||||| ||||||||||||||||||||
Db        320 GAGAAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGAAATTAATCCTGAAGTCTGG 262

Qy       2435 GCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAA 2494
             |||||||||||||||||||| || |||||||||||| || ||||||||||||||||||||
Db        261 GCAACAGAAGGACAATATGGACAAG-CAAAGAATGCCCATCTTGTTCAAGTTAAACTAAA 203

Qy       2495 GGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGG 2554
             |||||| || |||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        202 GGATTCRACCTCCTTTCCCTACCAAAAGCAGTACCCCCTTAGACCCAAGGCCCAACAAGG 143

Qy       2555 ATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCC 2614
               |||||||||||||||||| ||||||||||||||| ||||||||||||||  |  ||
Db        142 GCTCCAAAAGATTGTTAAGGACATAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCC 83

Qy       2615 CTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTA 2674
             ||||| || | |         |||   |  | ||||||||||   ||||||||| |||||
Db         82 CTGCAATACTCC-------AATTTTAGGAGTACAGAAACCCAACAGACAGTGGA-GGTTA 31

Qy       2675 GTGCAAGATCTCAGGATTATCAATG 2699
             |||||||||||||||||||||||||
Db         30 GTGCAAGATCTCAGGATTATCAATG 6


RESULT 139
US-09-925-065A-926128/c
; Sequence 926128, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926128
;    LENGTH: 735
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-926128

  Query Match             7.1%;  Score 532.6;  DB 1;  Length 735;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  660;  Conservative    1;  Mismatches   66;  Indels   18;  Gaps    8;

Qy       1958 AATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAAC 2017
             ||||||||||||||||||||||||||||||||  ||||||  | ||||||||||||||||||
Db        735 AATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTAAC 676

Qy       2018 CAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACC 2077
             |||||  ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        675 CAGATTTTCCCGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACC 616

Qy       2078 CTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGAC 2134
             |||||||||||| ||||| |||||||||||||||||||||||| | |   ||||||||||||
Db        615 CTCACAGAGCCCCGGGTACGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGGAC 556

Qy       2135 ACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTC 2194
             ||||| |  || |||||  |||||||||||  ||||||||||||||||||||||| |||||||||
Db        555 ACTGGAGTGGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCCTCTAGATCTGTC 496

Qy       2195 ACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAG 2254
             |||||  |   | | |  | || ||| | ||||||||||||||| ||:|||||||||| |
Db        495 ACTATCC---GAGCGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTGAC 440

Qy       2255 TTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCA 2314
             || || ||||| || ||||| ||||| |||||||||||||||||||| ||||||||||
Db        439 TTGTG-ACTGGAGAACTTTACTCTTTCCACATGCTTTTCTAATTATGCCTGAAAGCCCCA 381

Qy       2315 CTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAG 2374
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CTCCCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAG 321

Qy       2375 GAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGG 2434
             ||||||||||||| |||||||| |||||||||||||||| ||||||||||||||||||||
Db        320 GAGAAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGAAATTAATCCTGAAGTCTGG 262

Qy       2435 GCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAA 2494
             |||||||||||||||||||| || |||||||||||| || ||||||||||||||||||||
Db        261 GCAACAGAAGGACAATATGGACAAG-CAAAGAATGCCCATCTTGTTCAAGTTAAACTAAA 203

Qy       2495 GGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGG 2554
             |||||| || |||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        202 GGATTCRACCTCCTTTCCCTACCAAAAGCAGTACCCCCTTAGACCCAAGGCCCAACAAGG 143

Qy       2555 ATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCC 2614
               |||||||||||||||||| ||||||||||||||| ||||||||||||||  |  ||
Db        142 GCTCCAAAAGATTGTTAAGGACATAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCC 83

Qy       2615 CTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTA 2674
             ||||| || | |         |||   |  | ||||||||||   ||||||||| |||||
Db         82 CTGCAATACTCC-------AATTTTAGGAGTACAGAAACCCAACAGACAGTGGA-GGTTA 31

Qy       2675 GTGCAAGATCTCAGGATTATCAATG 2699
             |||||||||||||||||||||||||
Db         30 GTGCAAGATCTCAGGATTATCAATG 6


RESULT 140
US-08-979-847-8
; Sequence 8, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  645 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| ||||||| |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             | | |||  |||||||:||||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 141
US-10-430-442-8
; Sequence 8, Application US/10430442
; Publication No. US20030186391A1
;   GENERAL INFORMATION:
;        APPLICANT: Herve PERRON
;                   Frederic BESEME
;                   Frederic BEDIN
;                   Glaucia PARANHOS-BACCALA
;                   Florence KOMURIAN-PRADEL
;                   Colette JOLIVET
;                   Bernard MANDRAND
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 92
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Oliff & Berridge
;             STREET: 700 South Washington Street, Suite 300
;             CITY: Alexandria
;             STATE: Virginia
;             COUNTRY: U.S.A.
;             ZIP: 22314
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/430,442
;             FILING DATE: 07-May-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/691,563
;             FILING DATE: 02-AUG-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Berridge, William P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 38588
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 645 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-430-442-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
```

```
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| ||||||| |:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:||||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  |||||||||:||||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:|||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 142
US-10-114-104-8
; Sequence 8, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 645 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-114-104-8

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| ||||||| |:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:||||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  |||||||||:||||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593
```

```
Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              |||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 143
US-10-717-580-48
; Sequence 48, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-48

  Query Match             7.0%;  Score 522.6;  DB 1;  Length 530;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches  516;  Conservative   11;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 60

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 120

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 180

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||| ||||
Db        181 TGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 240

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 300

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        301 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 360

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              ||||||||||||||||||||||:|||||||||||||||||||||||||||:||||||||
Db        361 TCCATTCCTTGGAATCCGTGAGGCCAACGAACTCCAGGTCAGAGAATACGAAGCTTGCCA 420

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCT 551
              |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        421 CCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCT 480

Qy        552 GTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCCM 601
              ||||||||||||||||||:||||||||||:||||:::||||||||||:

Db        481 GTGAGCAAGGACCCCCCGGTGACATTTTGGCGACCACCAACGGACATCCC 530


RESULT 144
US-09-925-065A-928579
; Sequence 928579, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928579
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928579

  Query Match             7.0%;  Score 522.6;  DB 1;  Length 719;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  627;  Conservative    2;  Mismatches   82;  Indels   14;  Gaps    5;

Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
              |||||||||||||||||||| ||||||| |||||||||||||||||||| |||||||| |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60

Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
              |||| || |||||||||||| |||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                   |||||||||||||| |||||||||||  ||||||||||||   || ||||||||||
Db        121 TAAAAAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
              |||||||||||||  |||||||||  ||| |||  ||||||||||||||||||| ||||
Db        181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCCGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
              |||||||||||||||||||||||||||| | ||||| ||||| ||||| |||||||||| |
Db        241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
              ||| ||  ||| |||||||||||||||||||||||||| |||||| |||| | ||||||
Db        301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
              |||||||||||||||||||||||||||||||||||||||||||||  |||||||| |||
Db        361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
              |||||||||||| ||||||||||:| || |||||||||||||||||| |||||| | |
Db        421 CCCATGCCATCACTCTCACAGAGCSCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                |||||||||||||||  | | ||||| |||||||||| ||||||||||||||||||
Db        481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
```

```
                |||||||  ||||||||   | |  | |||  | || ||| |||||||||||||||| ||:
Db        541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
               ||||||||||||||| |||   ||| |  ||| |   |||||||||| |||| |||| |
Db        597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 2360
               |  |||||||||  || | ||||||||||||||||||||||||||||||||||||||||
Db        655 GTCTGAAAGCCCTGCTCCTTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 714

Qy       2361 ACACC 2365
               ||| |
Db        715 ACATC 719


RESULT 145
US-09-925-065A-928579
; Sequence 928579, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928579
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928579

  Query Match             7.0%;  Score 522.6;  DB 1;  Length 719;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  627;  Conservative    2;  Mismatches   82;  Indels   14;  Gaps    5;

Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
               |||||||||||||||||||| ||||||| |||||||||||||||||||| |||||||| |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60

Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
               |||| || |||||||||| |||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                    |||||||||||||| |||||||||| ||||||||||||   || ||||||||||
Db        121 TAAAAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
               |||||||||||||  |||||||||  ||| |||  |||||||||||||||||||| ||||
Db        181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCCGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
               |||||||||||||||||||||||||| | ||||| ||||| ||||| |||||||||| |
Db        241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
               ||| ||  ||| |||||||||||||||||||||||||||| |||||| |||| | |||||
Db        301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
               |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db        361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
               |||||||||||| |||||||||:| || |||||||||||||||||| |||||| | |
Db        421 CCCATGCCATCACTCTCACAGAGCSCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                 |||||||||||||||||  | | ||||| |||||||||| ||||||||||||||||
Db        481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
               ||||||| ||||||||   | |  | |||  | || ||| ||||||||||||||||| ||:
Db        541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
               |||||||||||||||| |||   ||| |  ||| |   |||||||||| |||| |||| |
Db        597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 2360
               |  |||||||||  || | ||||||||||||||||||||||||||||||||||||||||
Db        655 GTCTGAAAGCCCTGCTCCTTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 714

Qy       2361 ACACC 2365
               ||| |
Db        715 ACATC 719


RESULT 146
US-09-925-065A-928577
; Sequence 928577, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928577
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928577

  Query Match             7.0%;  Score 521.4;  DB 1;  Length 719;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  627;  Conservative    1;  Mismatches   83;  Indels   14;  Gaps    5;

Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
               |||||||||||||||||||| ||||||| |||||||||||||||||||| |||||||| |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60

Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
               |||| || |||||||||| |||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                    |||||||||||||| |||||||||| ||||||||||||   || ||||||||||
Db        121 TAARAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180
```

```
Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
              ||||||||||||||  |||||||||  ||| |||  ||||||||||||||||||| ||||
Db        181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCCGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
              |||||||||||||||||||||||||||| | ||||| ||||| ||||| ||||||||| |
Db        241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
              ||| ||  ||| |||||||||||||||||||||||||||| |||||| |||| | |||||
Db        301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
              |||||||||||||||||||||||||||||||||||||||||||||||  |||||||| |||
Db        361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
              |||||||||||||  ||||||||||  | ||  |||||||||||||||||  |||||| | |
Db        421 CCCATGCCATCACTCTCACAGAGCGCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                |||||||||||||||  | | |||||  ||||||||||  ||||||||||||||||
Db        481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
              |||||||  ||||||||   | |  | |||  | || ||| ||||||||||||||||  ||:
Db        541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
              |||||||||||||||  |||   ||| |  ||| |   |||||||||  |||| |||| |
Db        597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 2360
              |  |||||||||  || | ||||||||||||||||||||||||||||||||||||||||
Db        655 GTCTGAAAGCCCTGCTCCTTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 714

Qy       2361 ACACC 2365
              ||| |
Db        715 ACATC 719
```

```
RESULT 147
US-09-925-065A-928577
; Sequence 928577, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928577
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-928577

  Query Match             7.0%;  Score 521.4;  DB 1;  Length 719;
  Best Local Similarity   86.5%;  Pred. No. 0;
```

```
  Matches  627;  Conservative    1;  Mismatches   83;  Indels   14;  Gaps    5;

Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
              |||||||||||||||||||| ||||||| |||||||||||||||||| ||||||||  |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60

Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
              |||| || |||||||||| |||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                  |||||||||||||| |||||||||| ||||||||||||  || ||||||||||
Db         121 TAARAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
              |||||||||||||  |||||||||  ||| |||  ||||||||||||||||||| ||||
Db         181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCCGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
              |||||||||||||||||||||||||||| | ||||| ||||| ||||| ||||||||| |
Db         241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
              ||| ||  ||| |||||||||||||||||||||||||||| |||||| |||| | |||||
Db         301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
              |||||||||||||||||||||||||||||||||||||||||||||||  |||||||| |||
Db         361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
              |||||||||||||  |||||||||  | ||  |||||||||||||||||  |||||| | |
Db         421 CCCATGCCATCACTCTCACAGAGCGCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                |||||||||||||||  | | |||||  ||||||||||  ||||||||||||||||
Db         481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
              |||||||  ||||||||   | |  | |||  | || ||| ||||||||||||||||  ||:
Db         541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
              |||||||||||||||  |||   ||| |  ||| |   |||||||||  |||| |||| |
Db         597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 2360
              |  |||||||||  || | ||||||||||||||||||||||||||||||||||||||||
Db         655 GTCTGAAAGCCCTGCTCCTTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTAT 714

Qy       2361 ACACC 2365
              ||| |
Db         715 ACATC 719
```

```
RESULT 148
US-10-029-386-22647/c
; Sequence 22647, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22647 `
;   LENGTH: 591
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL137799.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 1.00e-35
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
US-10-029-386-22647

  Query Match             6.9%;  Score 520.3;  DB 1;  Length 591;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  523;  Conservative   40;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3440 TTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAG 3499
             ||:||:|||::||||||||||||||||:||:| ||||:|||||||||||||||:|||:||
Db        591 TTCTTTATATGTCACAGAAAAAAACAGGAATAACTCTAGGAGTCCTTACACAGGTCCAAG 532

Qy       3500 GGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             | |:||||||||||||::||||: ||||||:||||||||:||||||||||||||||||||
Db        531 GAACGAGCTTGCAACCCGTGGCAGACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 472

Qy       3560 GRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAA 3619
             |:|:||||||||||:||||| |||||||||||||| |::|||||||||||||||||||||
Db        471 GGCCTCATTGTTTATGGGTAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTTAAAA 412

Qy       3620 TAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTG 3679
             ||||||||| :|||||||||||||||||||||||||||||:||||::||||||||||||
Db        411 TAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTG 352

Qy       3680 CTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTG 3739
             |||||||||| |||||||||||||| |: ||||||||||: |||| ||||||||||||
Db        351 CTAAAGGAGACCTGTGGCTGTCAGACAGCCATTTACTTAAATACCAGGGTCTATTACTTG 292

Qy       3740 AARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAG 3799
             ||:||||||||||||:|||||||||||| |||||||||||||||:| ||||||||||||
Db        291 AAGGGCCAGTGCTGCAACTGTGCACTTGGGCAACTCTTAACCCAGCCACATTTCTTCCAG 232

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
             ||| |||| ||||||||:||: |||||||||||| :| ||| ||||||||||||| |||
Db        231 ACATTGAAAAAAAGATAGAACGTAACTGTCAACAGGTGATTGCTCAAACCTATGCCGCTC 172

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
             ||||  |||||:||||:||||:|||||||||||:|||| |||||||||||| |||||||
Db        171 GAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGACC-TCAACTTGTATATTGATGGA 113

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
             |||||||||||||||||| ||||||||||||: |||||||||||||||||||||||||||
Db        112 AGTTCCTTTGTAGAAAAAAGACTTCGAAAAGCAGGGTATGCAGTGGTCAGTGATAATGGA 53

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             |||:|||||| |||||||||||||||||||||||| |||:||||:||| |||
Db          52 ATACTTGAAAATAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAAAAC 1


RESULT 149
US-09-925-065A-935057
; Sequence 935057, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 935057
;   LENGTH: 629
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-935057

  Query Match             6.9%;  Score 519.3;  DB 1;  Length 629;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  541;  Conservative   33;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
             |||||||||||||||||:|||:| ||| |||||||| ||||||||:|||||||:|||||||
Db          1 GCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAATATA 60

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
             |||||||||||:|:| ||||:|||||||| |||||||||||::|||||||:|:||||
Db         61 TATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAA 120

Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
             |||||||||||||||||:|||:||||||||||||:|||||||||||||||||||:||||
Db        121 AGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTA 180

Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAA 4264
             |||:||||:|||||||||||||:||||||||:||| |||||||||:|||||||||||| |||
Db        181 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAGAAA 240

Qy       4265 GGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCA 4324
             ||||||:||||||||||:|||:||  :  |||||  :|||||||| :|||||||  ||
Db        241 GGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCCACA 300

Qy       4325 AGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTC 4384
             |||| |||||||||||||||||||||||||||    |   ||| |||||| |||||||||| ||
Db        301 AGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCCCTC 360

Qy       4385 CGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---C 4441
             :||||| |||||||||||||||||||||||||||||||||||| ||||| |||||| ||| ||||
Db        361 YGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAGGGAACTTCATGAGGACAT 420

Qy       4442 AGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACT 4501
             | ||  |||||||| || || ||||||| |  |||| ||||||||||||||||||| ||
Db        421 ACTTCCCTCCCCTCCAGATGGCTAGCCACCAATAAAGGAAAAATACTTTTGCCTGCAGCT 480

Qy       4502 ATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACC 4561
             | ||||| |||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db        481 AACCAATAGAAATTACTTAAAACCCTTCATCAAACCTTCCACTTAGGCATTGATAGCACC 540

Qy       4562 CATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATA 4621
             ||| |:|||||||||| |||||||||||||||||||||||||||||||||||||||:|||
Db        541 CATGAGATGGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATA 600

Qy       4622 KTCAGGGCCTGTGAAKTGTGCCARARAAA 4650
             :|||||||||||| ||:| |||||:|:|||
Db        601 GTCAGGGCCTGTAAAGTCTGCCAAAGAAA 629


RESULT 150
US-09-925-065A-935057
; Sequence 935057, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 935057
;   LENGTH: 629
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-935057

  Query Match             6.9%;  Score 519.3;  DB 1;  Length 629;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  541;  Conservative   33;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
             ||||||||||||||:|||:| ||| |||||||| ||||||||:||||||||:|||||||
Db          1 GCAGAACTAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAATATA 60

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
             ||||||||||||:|:| ||||:|||||||| ||||||||||||::|||||||:|:||||
Db         61 TATACAGACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAA 120

Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
             ||||||||||||||||||:|||:||||||||||||:|||||||||||||||||||:||||
Db        121 AGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTA 180

Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAA 4264
             |||:||||:|||||||||||||:|||||||:||| |||||||||:|||||||||| |||
Db        181 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAGAAA 240

Qy       4265 GGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCA 4324
             ||||||:||||||||||:|||:|| :  |||||  :|||||||| :||||||||  ||
Db        241 GGAAAGGAAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCCACA 300

Qy       4325 AGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTC 4384
             |||| |||||||||||||||||||||||||    |   ||| |||||| ||||||||||  ||
Db        301 AGGCTGGACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCCCTC 360

Qy       4385 CGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---C 4441
             :||||| |||||||||||||||||||||||||| ||||| |||||| ||| ||||
Db        361 YGGGAAGCCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAGGGAACTTCATGAGGACAT 420

Qy       4442 AGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACT 4501
             | ||  ||||||||  || || |||||||  |  |||| ||||||||||||||||||  ||
Db        421 ACTTCCCTCCCCTCCAGATGGCTAGCCACCAATAAAGGAAAAATACTTTTGCCTGCAGCT 480

Qy       4502 ATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACC 4561
             | ||||| |||||||||||||||||||||||||||||||| ||||||||||| ||||||||
Db        481 AACCAATAGAAATTACTTAAAACCCTTCATCAAACCTTCCACTTAGGCATTGATAGCACC 540

Qy       4562 CATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATA 4621
             ||| |:|||||||||| ||||||||||||||||||||||||||||||||||||||:|||
Db        541 CATGAGATGGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATA 600

Qy       4622 KTCAGGGCCTGTGAAKTGTGCCARARAAA 4650
             :||||||||||| ||:| |||||:|:|||
Db        601 GTCAGGGCCTGTAAAGTCTGCCAAAGAAA 629


RESULT 151
US-08-979-847-102
; Sequence 102, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  635 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-102

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
             ||||||||||||:|||||||||||||| ||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
             |||||||||||||:|||||| | || ||||||||||||||  |||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
             |||||||| |||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
             : || |||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||  |: |||||||:|||||||:|||||||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
```

```
Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              ||||||||| |||||| |||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | |||||||||||||||||||||||:||||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||| ||||||||||  ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 152
US-10-114-104-102
; Sequence 102, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 102:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 635 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 102:
```

US-10-114-104-102

Query Match 6.9%; Score 517.4; DB 1; Length 635;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;

```
Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
              ||||||||||||:|||||||||||||| ||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:||||||| | || |||||||||||||||  |||||||||||||| |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              ||||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              ||||||||||||||||||||||  |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              ||||||||| |||||| |||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | |||||||||||||||||||||||:||||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||| ||||||||||  ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 153
US-10-632-793-19
; Sequence 19, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
```

```
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-793-19

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:|||||||||||||||||||||||||| | ||||| || |||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             |||||||||||||||||||||||||||||  ||||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             |||||||||||||||:||||||||||||||||||:|||||||||   |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             |||||||||||||||| |||||| ||||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             |||||||||:|||||||||:|||||||||||||| |||:|||||||:|||||||  |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
               || |||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||| |||||||| ||||||||||||||||||||||||||:|||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||||||||||||| |||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 154
US-10-717-580-31
; Sequence 31, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Penv-C15 probe
US-10-717-580-31

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:|||||||||||||||||||||||||| | ||||| || |||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             |||||||||||||||||||||||||||||  ||||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             |||||||||||||||:||||||||||||||||||:|||||||||   |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             |||||||||||||||| |||||| ||||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             |||||||||:|||||||||:|||||||||||||| |||:|||||||:|||||||  |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
               || |||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||| |||||||| ||||||||||||||||||||||||||:|||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||||||||||||| |||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 155
US-09-925-065A-919228
; Sequence 919228, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
```

```
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919228
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919228

  Query Match             6.9%;  Score 513.6;  DB 1;  Length 598;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  522;  Conservative   41;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:|||||||||:|||||||||||||| :||:|||::|||||| |||||||||:||:||||
Db         15 TAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCT 73

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||||||:|||:|||||: ||| ||||||||::||||::||||||:||||
Db         74 CTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAG 133

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||: |||||||| ||||||||||
Db        134 GAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||| ||||||||| :|||||||||||||||||||||||||
Db        194 GTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCT 253

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||||:|||||
Db        254 CATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 313

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||| | |||||||:| ||||||||||:|||||||||||||||||||
Db        314 CTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACT 373

Qy       3775 CTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAA 3834
             ||||||||||:| |||||||||||| |||||||||||||||||:||:|||||||||||||||
Db        374 CTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAKAACATAACTGTCAACAA 433

Qy       3835 RTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCY 3894
             :|||||  |||||||||| |||||||| |||||||||||:||||:|||||: ||||||||||:
Db        434 GTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCC 493

Qy       3895 GACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGG 3954
             |||| |||||||||||||||||||||||||||| |||||||||| |||| |||||:|||
Db        494 GACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGG 552

Qy       3955 GTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
             ||||||||| |||||||||||||||||:||||||||||||||||||
Db        553 GTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTC 598


RESULT 156
US-09-925-065A-919228
; Sequence 919228, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919228
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919228

  Query Match             6.9%;  Score 513.6;  DB 1;  Length 598;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  522;  Conservative   41;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:|||||||||:|||||||||||||| :||:|||::|||||| |||||||||:||:||||
Db         15 TAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCT 73

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||||||:|||:|||||: ||| ||||||||::||||::||||||:||||
Db         74 CTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAG 133

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||: |||||||| ||||||||||
Db        134 GAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||| ||||||||| :|||||||||||||||||||||||||
Db        194 GTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCT 253

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||||:|||||
Db        254 CATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 313

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||| | |||||||:| ||||||||||:|||||||||||||||||||
Db        314 CTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACT 373

Qy       3775 CTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAA 3834
             ||||||||||:| |||||||||||| |||||||||||||||||:||:|||||||||||||||
Db        374 CTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAKAACATAACTGTCAACAA 433

Qy       3835 RTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCY 3894
             :|||||  |||||||||| |||||||| |||||||||||:||||:|||||: ||||||||||:
Db        434 GTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCC 493

Qy       3895 GACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGG 3954
             |||| |||||||||||||||||||||||||||| |||||||||| |||| |||||:|||
Db        494 GACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGG 552

Qy       3955 GTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
             ||||||||| |||||||||||||||||:||||||||||||||||||
Db        553 GTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTC 598


RESULT 157
US-09-925-065A-212368
; Sequence 212368, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 212368
;   LENGTH: 651
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-212368

  Query Match             6.8%;  Score 512.8;  DB 1;  Length 651;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  555;  Conservative   57;  Mismatches   39;  Indels    6;  Gaps    5;

Qy       2841 TTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCC 2900
             || |||||||||||||| |||| ||||||:  |||||||||||||||||||:||||||||
Db          1 TTATTTGCCTTTGAAGATCCTTCGAACCCAATGTCTCAACTCACCTGGACTGTTTTACCC 60

Qy       2901 CAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCAR 2960
             || |||||||| |||||:|||||||||||||||||||||||||||||||||||||||:||:
Db         61 CAGGGGTTCAGAGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAA 120

Qy       2961 TYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRT 3020
             |::|||||||||||||||||||||||| :|||  |||||||||||||||:||: ||::|
Db        121 TTCTCATACCTGGACACTCTTGTCCTTTGGTACATGGATGATTTACTTTTAGCCACCCAT 180

Qy       3021 TCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGC 3080
             |||||| |||||||||||||||||||||||:|||||:||||||||||:|||||||||
Db        181 TCAGAAGCCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGC 240

Qy       3081 TACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTA 3140
             ||||: | ||||||||:|:|:|||||:|||||||||||| || ||||||||||||:|||
Db        241 TACAA-GATTCCAAACCAAAGGCTCAGCTCTGCTCACAGAAGTTTAAATACTTAGGGCTA 299

Qy       3141 ARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTAT 3200
             |:||||||:||||| ||||:|||||||||| ||||:::|| |||||||||||||||| |
Db        300 AAATTATCSAAAGGTACCAGGGCCCTCAGTGGGGAACATATGCAGCCTATACTGGCTTGT 359

Qy       3201 CCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGA 3260
             |||||||:||| ||||| |||||||||||:|::|| ||||||:|||:|||:|||||||| |
Db        360 CCTCATCCCAAGACCCT-AAGCAACTAAGAGGGTT-CTTGGCATAACAGGTTTCTGCCAA 417

Qy       3261 AWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAAC 3320
             |:||||||| |||||||:||||:|||||||||| :|||||:|||||:|||||||||||||
Db        418 ATATGGATT-CCCAGGTATGGCAAAATAGCCAGATCATTATATACACTAATTAAGGAAAC 476

Qy       3321 TCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGC 3380
             |||||||||||||||||||||:|||| ||||:|:|||||:: |||  ||||||||||
Db        477 TCAGAAAGCCAATACCCATTTAGTAAGGTGGACACCTGAAGCACAAGCAGCTTTCCAGGC 536

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
               |   ||  | | ||  ||||||||: |||||||| |:|||||| :||||||| |||
Db        537 TC--TAAAGAAGGCTCTAACCCAAGTTCCAGTGTTACGCTTGCCAAAGGGGCAAGATTTT 594

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCR 3497
             |:||:| |::|||||||||||||||:||:||||||:||||||||||||||:|||:
Db        595 TCTTTAAATGTCACAGAAAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCG 651


RESULT 158
US-09-925-065A-212368
; Sequence 212368, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 212368
;   LENGTH: 651
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-212368

  Query Match             6.8%;  Score 512.8;  DB 1;  Length 651;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  555;  Conservative   57;  Mismatches   39;  Indels    6;  Gaps    5;

Qy       2841 TTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCC 2900
             || |||||||||||||| |||| ||||||:  |||||||||||||||||||:||||||||
Db          1 TTATTTGCCTTTGAAGATCCTTCGAACCCAATGTCTCAACTCACCTGGACTGTTTTACCC 60

Qy       2901 CAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCAR 2960
             || |||||||| |||||:|||||||||||||||||||||||||||||||||||||||:||:
Db         61 CAGGGGTTCAGAGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAA 120

Qy       2961 TYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRT 3020
             |::|||||||||||||||||||||||| :|||  |||||||||||||||:||: ||::|
Db        121 TTCTCATACCTGGACACTCTTGTCCTTTGGTACATGGATGATTTACTTTTAGCCACCCAT 180

Qy       3021 TCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGC 3080
             |||||| |||||||||||||||||||||||:|||||:||||||||||:|||||||||
Db        181 TCAGAAGCCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGC 240

Qy       3081 TACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTA 3140
             ||||: | ||||||||:|:|:|||||:|||||||||||| || ||||||||||||:|||
Db        241 TACAA-GATTCCAAACCAAAGGCTCAGCTCTGCTCACAGAAGTTTAAATACTTAGGGCTA 299

Qy       3141 ARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTAT 3200
             |:||||||:||||| ||||:|||||||||| ||||:::|| |||||||||||||||| |
Db        300 AAATTATCSAAAGGTACCAGGGCCCTCAGTGGGGAACATATGCAGCCTATACTGGCTTGT 359

Qy       3201 CCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGA 3260
             |||||||:||| ||||| |||||||||||:|::|| ||||||:|||:|||:|||||||| |
Db        360 CCTCATCCCAAGACCCT-AAGCAACTAAGAGGGTT-CTTGGCATAACAGGTTTCTGCCAA 417

Qy       3261 AWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAAC 3320
             |:||||||| |||||||:||||:|||||||||| :|||||:|||||:|||||||||||||
Db        418 ATATGGATT-CCCAGGTATGGCAAAATAGCCAGATCATTATATACACTAATTAAGGAAAC 476

Qy       3321 TCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGC 3380
             |||||||||||||||||||||:|||| ||||:|:|||||:: |||  ||||||||||
Db        477 TCAGAAAGCCAATACCCATTTAGTAAGGTGGACACCTGAAGCACAAGCAGCTTTCCAGGC 536

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
               |   ||  | | ||  ||||||||: |||||||| |:|||||| :||||||| |||
Db        537 TC--TAAAGAAGGCTCTAACCCAAGTTCCAGTGTTACGCTTGCCAAAGGGGCAAGATTTT 594

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCR 3497
             |:||:| |::|||||||||||||||:||:||||||:||||||||||||||:|||:
Db        595 TCTTTAAATGTCACAGAAAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCG 651


RESULT 159
US-09-925-065A-896106
; Sequence 896106, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
```

```
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 896106
;   LENGTH: 619
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-896106

  Query Match             6.8%;  Score 510.9;  DB 1;  Length 619;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  535;  Conservative   29;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4018 TYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGY 4077
                | ||||:||  |||||||||||:|||:| ||||||||||||||||||||: |||||||:
Db          1 TGAGCTGGCCAAACTAATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAGAAAAGGGT 60

Qy       4078 AAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATG 4137
                |||||||||||||||| |:|:|||||:|||||||| | ||||||||||::||||||||
Db         61 AAATATATATACAGACTATAAGTATGCTTACCTAGTCCTTCATGCCCATGCAGCAATATG 120

Qy       4138 SARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAG 4197
                :|:|||||||||||||||||||: | :||||||||||||:|||||||||||||||||||
Db        121 GAGAGAAAGGGAATTCCTAACTTCCAAAGGAACACCTATCAAACATCAGGAAGCCATTAG 180

Qy       4198 GARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCA 4257
                || |||||||:|||  |||||||||||||:||||||||:|||| ||||||||:|||||||
Db        181 GATATTATTATTGGTGGTACAGAAACCTAAAGAGGTGGCAGTCCTACACTGCTGGGGTCA 240

Qy       4258 TCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAG 4317
                ||| |||  ||||:|||||||||||:|||:|||:: |||||||:|||||||||||:|||||
Db        241 TCAGAAAAAAAAGGAAAGGGAAATAGAAGGGAACYACCAAGCAGATATTGAAGCCAAAAG 300

Qy       4318 AGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAA 4377
                ||| |||||||||||||||||||||||||||||||  ||   ||| |||| | ||||||
Db        301 AGCCGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAA 360

Qy       4378 TCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACG 4437
                 ||| ||| ||||| ||| |||||||||||||||||||||| | ||||| ||||||||||
Db        361 CCCCCTCCAGGAAAGCAATCCCCAGTACTCAGCAGGAGAAATAAAATGGAGAACCTCACG 420

Qy       4438 AGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCC 4494
                |||    | ||| |||||||| ||| || |||||||  |||||||  ||||| ||||||||
Db        421 AGGACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATGCTTTTGCC 480

Qy       4495 TGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGA 4554
                |||| ||| |||||||||||||||||||||||||||||| ||||||||||||||||||| ||
Db        481 TGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGA 540

Qy       4555 TAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 4614
                ||||||||||||:|||||||||| |||||||||||||| ||||||||||||||||||||
Db        541 TAGCACCCATCAGATGGCCAAATTATTATTTACTGGATCAGGCCTTTTCAAAACTATCAA 600

Qy       4615 GCARATAKTCAGGGCCTGT 4633
                |||: ||:||||||||||||
Db        601 GCAGGTAGTCAGGGCCTGT 619


RESULT 160
US-09-925-065A-896106
```

```
; Sequence 896106, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 896106
;   LENGTH: 619
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-896106

  Query Match             6.8%;  Score 510.9;  DB 1;  Length 619;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  535;  Conservative   29;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4018 TYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGY 4077
                | ||||:||  |||||||||||:|||:| ||||||||||||||||||||: |||||||:
Db          1 TGAGCTGGCCAAACTAATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAGAAAAGGGT 60

Qy       4078 AAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATG 4137
                |||||||||||||||| |:|:|||||:|||||||| | ||||||||||::||||||||
Db         61 AAATATATATACAGACTATAAGTATGCTTACCTAGTCCTTCATGCCCATGCAGCAATATG 120

Qy       4138 SARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAG 4197
                :|:|||||||||||||||||||: | :||||||||||||:|||||||||||||||||||
Db        121 GAGAGAAAGGGAATTCCTAACTTCCAAAGGAACACCTATCAAACATCAGGAAGCCATTAG 180

Qy       4198 GARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCA 4257
                || |||||||:|||  |||||||||||||:||||||||:|||| ||||||||:|||||||
Db        181 GATATTATTATTGGTGGTACAGAAACCTAAAGAGGTGGCAGTCCTACACTGCTGGGGTCA 240

Qy       4258 TCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAG 4317
                ||| |||  ||||:|||||||||||:|||:|||:: |||||||:|||||||||||:|||||
Db        241 TCAGAAAAAAAAGGAAAGGGAAATAGAAGGGAACYACCAAGCAGATATTGAAGCCAAAAG 300

Qy       4318 AGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAA 4377
                ||| |||||||||||||||||||||||||||||||  ||   ||| |||| | ||||||
Db        301 AGCCGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAA 360

Qy       4378 TCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACG 4437
                 ||| ||| ||||| ||| |||||||||||||||||||||| | ||||| ||||||||||
Db        361 CCCCCTCCAGGAAAGCAATCCCCAGTACTCAGCAGGAGAAATAAAATGGAGAACCTCACG 420

Qy       4438 AGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCC 4494
                |||    | ||| |||||||| ||| || |||||||  |||||||  ||||| ||||||||
Db        421 AGGACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATGCTTTTGCC 480

Qy       4495 TGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGA 4554
                |||| ||| |||||||||||||||||||||||||||||| ||||||||||||||||||| ||
Db        481 TGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGA 540

Qy       4555 TAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 4614
                ||||||||||||:|||||||||| |||||||||||||| ||||||||||||||||||||
Db        541 TAGCACCCATCAGATGGCCAAATTATTATTTACTGGATCAGGCCTTTTCAAAACTATCAA 600

Qy       4615 GCARATAKTCAGGGCCTGT 4633
                |||: ||:||||||||||||
```

```
Db        601 GCAGGTAGTCAGGGCCTGT 619


RESULT 161
US-09-925-065A-943242/c
; Sequence 943242, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 943242
;   LENGTH: 656
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-943242

  Query Match             6.8%;  Score 510.6;  DB 1;  Length 656;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  587;  Conservative    3;  Mismatches   58;  Indels    8;  Gaps    4;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             ||||||||||||||||||||| |||||||||||| |||||||||||||| |||||||||
Db        651 AAAAAACTTCAAAAGTCTGCCTTAGGCCCGGAGCGAAACTTAGAAACCCTGTTGAACTTG 592

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             |||||:|| |||||||||||||||||||||||||||||||   |||  |||||||| ||||
Db        591 GCAACCTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTAGAATGGGACAAACAGGAT 532

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |  |||||||||||||||||| ||||| ||| |||||||| ||||||||||||||||||
Db        531 TTTAAAAAAGGCCACCGCTTTACTCATGGCCCACAGGCAAGCGGACTTTGGAGGCTCTGG 472

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             |||||||||||||||||||| |||| ||||||||||||||||||||||| |||||||||
Db        471 AAAAGGGAAAAGCTGGGCAAATCGAATACCTAATAGGGCTTGCTTCCAGTGTGGTCTACA 412

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||||||||||||||| ||||| || ||||||||||||
Db        411 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTCTTCCATGCCCCTT 352

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||:|||||||||||||||||||||||||||||||  ||| || ||||||||||||
Db        351 ATGTTAAGRGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGTTCCTTTGAGTCAGA 292

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             || |||||||||||||| |||||||||||||||||||||||||||| ||||||||| |||| ||||
Db        291 AGTCACTAACCAGATGACCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCACCAGCCCA 232

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
             |||||||||||||||||||||| ||||||||||||||  |||||||||| | |    ||
Db        231 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATCAAGGGCCAGGAGGTTAACTG 172

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
             |||||||||||||||| | ||||| |||||||||| |||||| |||||||||| ||
Db        171 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCATC 112

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
             |||||||||||||||| |   ||| |    || ||| | |||||||||||||| ||:||||
Db        111 CAGATCTGTCACTATCCAAGGGGTG---CTAGGACAGCCAGTCACTAGATAC-TTCTCCC 56

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
             |||||||||| ||| |||||||| ||||| | |||||| ||||||||||||||||
Db         55 AGCCACTAAGTCATG-GCTGGGGAACTTTACTTTTTTCATATGCTTTTCTAATTAT 1


RESULT 162
US-09-925-065A-943242/c
; Sequence 943242, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 943242
;   LENGTH: 656
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-943242

  Query Match             6.8%;  Score 510.6;  DB 1;  Length 656;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  587;  Conservative    3;  Mismatches   58;  Indels    8;  Gaps    4;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||||| |||||||||||| |||||||||||||| |||||||||
Db        651 AAAAAACTTCAAAAGTCTGCCTTAGGCCCGGAGCGAAACTTAGAAACCCTGTTGAACTTG 592

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             |||||:|| |||||||||||||||||||||||||||||||   |||  |||||||| ||||
Db        591 GCAACCTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTAGAATGGGACAAACAGGAT 532

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |  |||||||||||||||||| ||||| ||| |||||||| ||||||||||||||||||
Db        531 TTTAAAAAAGGCCACCGCTTTACTCATGGCCCACAGGCAAGCGGACTTTGGAGGCTCTGG 472

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             |||||||||||||||||||| |||| ||||||||||||||||||||||| |||||||||
Db        471 AAAAGGGAAAAGCTGGGCAAATCGAATACCTAATAGGGCTTGCTTCCAGTGTGGTCTACA 412

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||||||||||||||| ||||| || ||||||||||||
Db        411 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTCTTCCATGCCCCTT 352

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||:|||||||||||||||||||||||||||||||  ||| || ||||||||||||
Db        351 ATGTTAAGRGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGTTCCTTTGAGTCAGA 292

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             || |||||||||||||| |||||||||||||||||||||||||||| ||||||||| |||| ||||
Db        291 AGTCACTAACCAGATGACCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCACCAGCCCA 232

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
             |||||||||||||||||||||| ||||||||||||||  |||||||||| | |    ||
Db        231 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATCAAGGGCCAGGAGGTTAACTG 172
```

```
Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
             |||||||||||||||||| | ||||| |||||||||| |||||| |||||||||| ||
Db        171 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCATC 112

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
             |||||||||||||||| |   ||| |    || ||| | |||||||||||||| ||:||||
Db        111 CAGATCTGTCACTATCCAAGGGGTG---CTAGGACAGCCAGTCACTAGATAC-TTCTCCC 56

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
             |||||||||| |||  ||||||| ||||| | |||||| ||||||||||||||||
Db          55 AGCCACTAAGTCATG-GCTGGGGAACTTTACTTTTTTCATATGCTTTTCTAATTAT 1
```

RESULT 163
US-09-925-065A-923210
; Sequence 923210, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923210
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-923210

Query Match 6.8%; Score 510.6; DB 1; Length 657;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 587; Conservative 3; Mismatches 58; Indels 8; Gaps 4;

```
Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||||| |||||||||||| |||||||||||||||| |||||||||
Db          7 AAAAAACTTCAAAAGTCTGCCTTAGGCCCGGAGCGAAACTTAGAAACCCTGTTGAACTTG 66

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             |||||:|| |||||||||||||||||||||||||||||  |||  |||||||| ||||
Db         67 GCAACCTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTAGAATGGGACAAACAGGAT 126

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |  ||||||||||||||||||| ||||| ||| |||||||| ||||||||||||||||||
Db        127 TTTAAAAAAGGCCACCGCTTTACTCATGGCCCACAGGCAAGCGGACTTTGGAGGCTCTGG 186

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             |||||||||||||||||||| |||| |||||||||||||||||||| ||||||||
Db        187 AAAAGGGAAAAGCTGGGCAAATCGAATACCTAATAGGGCTTGCTTCCAGTGTGGTCTACA 246

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||||||||||||||||||| ||||| || ||||||||||||
Db        247 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTCTTCCATGCCCCTT 306

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||:|||||||||||||||||||||||||||||||||  ||| || ||||||||||
Db        307 ATGTTAAGRGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGTTCCTTTGAGTCAGA 366

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             || |||||||||||||| |||||||||||||||||||| |||||||||| ||| ||||
Db        367 AGTCACTAACCAGATGACCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCACCAGCCCA 426
```

```
Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
             |||||||||||||||||||||||| ||||||||||||||  |||||||||| | |   ||
Db        427 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATCAAGGGCCAGGAGGTTAACTG 486

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
             |||||||||||||||||| | ||||| |||||||||| |||||| |||||||||| ||
Db        487 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCATC 546

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
             |||||||||||||||| |   ||| |    || ||| | |||||||||||||| ||:||||
Db        547 CAGATCTGTCACTATCCAAGGGGTG---CTAGGACAGCCAGTCACTAGATAC-TTCTCCC 602

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
             |||||||||| |||  ||||||| ||||| | |||||| ||||||||||||||||
Db        603 AGCCACTAAGTCATG-GCTGGGGAACTTTACTTTTTTCATATGCTTTTCTAATTAT 657
```

RESULT 164
US-09-925-065A-923210
; Sequence 923210, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923210
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-923210

Query Match 6.8%; Score 510.6; DB 1; Length 657;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 587; Conservative 3; Mismatches 58; Indels 8; Gaps 4;

```
Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||||| |||||||||||| |||||||||||||||| |||||||||
Db          7 AAAAAACTTCAAAAGTCTGCCTTAGGCCCGGAGCGAAACTTAGAAACCCTGTTGAACTTG 66

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             |||||:|| |||||||||||||||||||||||||||||  |||  |||||||| ||||
Db         67 GCAACCTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTAGAATGGGACAAACAGGAT 126

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |  ||||||||||||||||||| ||||| ||| |||||||| ||||||||||||||||||
Db        127 TTTAAAAAAGGCCACCGCTTTACTCATGGCCCACAGGCAAGCGGACTTTGGAGGCTCTGG 186

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             |||||||||||||||||||| |||| |||||||||||||||||||| ||||||||
Db        187 AAAAGGGAAAAGCTGGGCAAATCGAATACCTAATAGGGCTTGCTTCCAGTGTGGTCTACA 246

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||||||||||||||||||| ||||| || ||||||||||||
Db        247 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTCTTCCATGCCCCTT 306

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||:|||||||||||||||||||||||||||||||||  ||| || ||||||||||
```

```
Db        307 ATGTTAAGRGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGTTCCTTTGAGTCAGA 366

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              || |||||||||||||| |||||||||||||||||||| |||||||||| ||| ||||
Db        367 AGTCACTAACCAGATGACCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCACCAGCCCA 426

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
              |||||||||||||||||||| ||||||||||||||  |||||||||| | |   ||
Db        427 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATCAAGGGCCAGGAGGTTAACTG 486

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
              |||||||||||||||| | ||||| |||||||||| |||||| |||||||||| ||
Db        487 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCATC 546

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
              |||||||||||||| |   ||| |    || ||| | |||||||||||||| ||:||||
Db        547 CAGATCTGTCACTATCCAAGGGGTG---CTAGGACAGCCAGTCACTAGATAC-TTCTCCC 602

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
              |||||||||| |||  ||||||| ||||| | |||||| ||||||||||||||||
Db        603 AGCCACTAAGTCATG-GCTGGGGAACTTTACTTTTTTCATATGCTTTTCTAATTAT 657
```

RESULT 165
US-10-040-916-48
; Sequence 48, Application US/10040916
; Publication No. US20020146769A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; LaVallie, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Evans, Cheryl
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,916
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,029
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: 08/686,878
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-040-916-48

Query Match 6.8%; Score 510.3; DB 1; Length 542;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 516; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

```
Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
              ||||| || |||||||||||||||||||| ||||| |||| |||||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
              ||||::|||||||||||||||||||||||| |||:|| |||||||| |||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
              | ||||| ||||| |||||||||||||||| |||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
              |||||||||||||||||||||||||||||||||||| ||||||||:|||||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
              ||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
              ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
              |||
Db        540 CAT 542
```

RESULT 166
US-09-925-065A-920958/c
; Sequence 920958, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 920958
; LENGTH: 626

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-920958

  Query Match             6.8%;  Score 508;  DB 1;  Length 626;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  532;  Conservative   53;  Mismatches   39;  Indels    6;  Gaps    3;

Qy       3171 GAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGR 3230
             | |  |:::|||||||||||||||||||||||||||| |:||||||||||||||||||||:
Db        626 GTGAGATGTATCCAGCCTATACTGGCTTATCCTCACCCCAAAACCCTAAAGCAACTAAGA 567

Qy       3231 GRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGC 3290
              ::|||||||||:|||:|||:||||||||||:|||||||  ||||||:  | :|||||||
Db        566 CGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-TCCAGGTACAGTGAAATAGC 508

Qy       3291 CAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATG 3350
             ||| :|||||:|||||:|||||||||||||||||||||||||||| |||||||:|||||||
Db        507 CAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATATCCATTTAGTAAGATG 448

Qy       3351 GAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCA 3410
             ||:|:|||||  :::|||  ||||||||||||  ||  | | |||   | |||||:||||
Db        447 GATACCTGAAACAGAAGCAGCTTTCCAGGCCCTAAAGAAGGCCCT---AACCAAGCCCCA 391

Qy       3411 GTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRT--CACAGAAAAAAACAGRA 3468
             ||||||||:|||||||  :|||||||||||||:||:|||::|  ||||  ||||||||||:|
Db        390 GTGTTAAGCTTGCCAATGGGGCAAGACTTTTYTTTATATGTCACACACAAAAAAACAGGA 331

Qy       3469 AYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTG 3528
             |:|||||||: |||||||||||||||:|||:|| ||:||| ||||||||::||||::|||||
Db        330 ATAGCTCTAAGAGTCCTTACACAGGTCCAAGAGATGAGTTTGCAACCCGTGGCATACCTG 271

Qy       3529 ASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCA 3588
             |:|||||||:|||||  ||||||||||||||:|:||||||||||: |||||||  ||||
Db        270 AGTAAGGAAACTGATGCAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGACGGCA 211

Qy       3589 GTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGG 3648
             ||| ||||::|||||||||||||||||||||||||||||:|||||| ||||||||||||
Db        210 GTAACAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATTTTACTGTGTGG 151

Qy       3649 ACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAAC 3708
             |||||||||||:||||||::|||||||||||||||||||||||||||||| |||||||||
Db        150 ACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCCGTCAGACAAA 91

Qy       3709 YGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGT 3768
              |||||||||||:|||||||||||||||||||:||||||||||||:|||| ||||||||
Db         90 AGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGT 31

Qy       3769 GCAACTCTTAACCCAGYCNCATTTCTTCCA 3798
             ||||||||||||||| | |||||||| ||
Db         30 GCAACTCTTAACCCAGGCACATTTCTTTCA 1


RESULT 167
US-09-925-065A-920958/c
; Sequence 920958, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 920958
;   LENGTH: 626
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-920958

  Query Match             6.8%;  Score 508;  DB 1;  Length 626;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  532;  Conservative   53;  Mismatches   39;  Indels    6;  Gaps    3;

Qy       3171 GAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGR 3230
             | |  |:::|||||||||||||||||||||||||||| |:||||||||||||||||||||:
Db        626 GTGAGATGTATCCAGCCTATACTGGCTTATCCTCACCCCAAAACCCTAAAGCAACTAAGA 567

Qy       3231 GRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGC 3290
              ::|||||||||:|||:|||:||||||||||:|||||||  ||||||:  | :|||||||
Db        566 CGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-TCCAGGTACAGTGAAATAGC 508

Qy       3291 CAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATG 3350
             ||| :|||||:|||||:|||||||||||||||||||||||||||| |||||||:|||||||
Db        507 CAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATATCCATTTAGTAAGATG 448

Qy       3351 GAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCA 3410
             ||:|:|||||  :::|||  ||||||||||||  ||  | | |||   | |||||:||||
Db        447 GATACCTGAAACAGAAGCAGCTTTCCAGGCCCTAAAGAAGGCCCT---AACCAAGCCCCA 391

Qy       3411 GTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRT--CACAGAAAAAAACAGRA 3468
             ||||||||:|||||||  :|||||||||||||:||:|||::|  ||||  ||||||||||:|
Db        390 GTGTTAAGCTTGCCAATGGGGCAAGACTTTTYTTTATATGTCACACACAAAAAAACAGGA 331

Qy       3469 AYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTG 3528
             |:|||||||: |||||||||||||||:|||:|| ||:||| ||||||||::||||::|||||
Db        330 ATAGCTCTAAGAGTCCTTACACAGGTCCAAGAGATGAGTTTGCAACCCGTGGCATACCTG 271

Qy       3529 ASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCA 3588
             |:|||||||:|||||  ||||||||||||||:|:||||||||||: |||||||  ||||
Db        270 AGTAAGGAAACTGATGCAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGACGGCA 211

Qy       3589 GTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGG 3648
             ||| ||||::|||||||||||||||||||||||||||||:|||||| ||||||||||||
Db        210 GTAACAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATTTTACTGTGTGG 151

Qy       3649 ACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAAC 3708
             |||||||||||:||||||::|||||||||||||||||||||||||||||| |||||||||
Db        150 ACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCCGTCAGACAAA 91

Qy       3709 YGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGT 3768
              |||||||||||:|||||||||||||||||||:||||||||||||:|||| ||||||||
Db         90 AGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGT 31

Qy       3769 GCAACTCTTAACCCAGYCNCATTTCTTCCA 3798
             ||||||||||||||| | |||||||| ||
Db         30 GCAACTCTTAACCCAGGCACATTTCTTTCA 1


RESULT 168
US-09-925-065A-210294/c
; Sequence 210294, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210294
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210294

  Query Match             6.8%;  Score 507.1;  DB 1;  Length 627;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  565;  Conservative    9;  Mismatches   48;  Indels    9;  Gaps    3;

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||
Db        627 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 568

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||||  |
Db        567 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 508

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        507 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 448

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        447 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 389

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             ||||||||||||||||||||||||    ||  | || ||||||||||||||||||||||||||
Db        388 TCTACACTGAACAAGATTTATACA---ATCACGTACCTAAGCCCCGCAACAAAAGAGTAC 332

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||| : ||||||||||||||||||||
Db        331 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCRGACTAGGTACTGGCATTGGCA 272

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db        271 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 212

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        211 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 152

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:|||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        151 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 92

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         91 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 32

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         31 AAGAAATTTGAGATTGATTACAACATAGAAC 1


RESULT 169
US-09-925-065A-210294/c
; Sequence 210294, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210294
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210294

  Query Match             6.8%;  Score 507.1;  DB 1;  Length 627;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  565;  Conservative    9;  Mismatches   48;  Indels    9;  Gaps    3;

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||
Db        627 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 568

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||||  |
Db        567 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 508

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        507 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 448

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        447 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 389

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             ||||||||||||||||||||||||    ||  | || ||||||||||||||||||||||||||
Db        388 TCTACACTGAACAAGATTTATACA---ATCACGTACCTAAGCCCCGCAACAAAAGAGTAC 332

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||| : ||||||||||||||||||||
Db        331 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCRGACTAGGTACTGGCATTGGCA 272

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db        271 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 212

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        211 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 152

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:|||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        151 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 92

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         91 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 32

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         31 AAGAAATTTGAGATTGATTACAACATAGAAC 1


RESULT 170
US-09-925-065A-210296/c
; Sequence 210296, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210296
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210296

  Query Match             6.8%;  Score 507.1;  DB 1;  Length 627;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  565;  Conservative    9;  Mismatches   48;  Indels    9;  Gaps    3;

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             |||| | |||||||||||:||| | ||||||| || |||||| |||||||||||||||||
Db        627 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 568

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||| | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        567 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 508

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             ||||||||||||||||||||||||     ||||||||||||||||||||||||||||||||
Db        507 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 448

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        447 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 389

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||:| |   | || ||||||||||||||||||||||||
Db        388 TCTACACTGAACAAGATTTATACARTCA---CGTACCTAAGCCCCGCAACAAAAGAGTAC 332

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| ||||||||||||||||   ||||||||||||||||||||||
Db        331 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCA 272

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||| |||||||
Db        271 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 212

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        211 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 152

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:|||||  |||||||||||||||:|||||| ||||:|| |||||||  ||||||||
Db        151 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 92

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         91 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 32

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         31 AAGAAATTTGAGATTGATTACAACATAGAAC 1
```

```
RESULT 171
US-09-925-065A-210296/c
; Sequence 210296, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210296
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210296

  Query Match             6.8%;  Score 507.1;  DB 1;  Length 627;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  565;  Conservative    9;  Mismatches   48;  Indels    9;  Gaps    3;

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             |||| | |||||||||||:||| | ||||||| || |||||| |||||||||||||||||
Db        627 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 568

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||| | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        567 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 508

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             ||||||||||||||||||||||||     ||||||||||||||||||||||||||||||||
Db        507 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 448

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        447 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 389

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||||:| |   | || ||||||||||||||||||||||||
Db        388 TCTACACTGAACAAGATTTATACARTCA---CGTACCTAAGCCCCGCAACAAAAGAGTAC 332

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| ||||||||||||||||   ||||||||||||||||||||||
Db        331 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCA 272

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||| |||||||
Db        271 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 212

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        211 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 152

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:|||||  |||||||||||||||:|||||| ||||:|| |||||||  ||||||||
Db        151 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 92

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         91 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 32
```

```
Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         31 AAGAAATTTGAGATTGATTACAACATAGAAC 1
```

RESULT 172
US-10-301-480-297639/c
; Sequence 297639, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 297639
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-297639

Query Match 6.8%; Score 507.1; DB 1; Length 628;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 565; Conservative 9; Mismatches 48; Indels 9; Gaps 3;

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||||
Db        628 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 569

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        568 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 509

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        508 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 449

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        448 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 390

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||    ||  | || ||||||||||||||||||||||||
Db        389 TCTACACTGAACAAGATTTATACA---ATCACGTACCTAAGCCCCGCAACAAAAGAGTAC 333

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||| : ||||||||||||||||||||
Db        332 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCRGACTAGGTACTGGCATTGGCA 273

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             ||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db        272 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 213

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 153

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        152 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 93

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         92 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 33

```
Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         32 AAGAAATTTGAGATTGATTACAACATAGAAC 2
```

RESULT 173
US-10-301-480-297641/c
; Sequence 297641, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 297641
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-297641

Query Match 6.8%; Score 507.1; DB 1; Length 628;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 565; Conservative 9; Mismatches 48; Indels 9; Gaps 3;

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||||
Db        628 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 569

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        568 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 509

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        508 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 449

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        448 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 390

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||:| |    | || ||||||||||||||||||||||||
Db        389 TCTACACTGAACAAGATTTATACARTCA---CGTACCTAAGCCCCGCAACAAAAGAGTAC 333

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||||    ||||||||||||||||||||
Db        332 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCA 273

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             ||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db        272 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 213

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 153

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        152 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 93

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         92 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 33

```
Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
```

```
Db         32 AAGAAATTTGAGATTGATTACAACATAGAAC 2
```

RESULT 174
US-10-301-480-911048/c
; Sequence 911048, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911048
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-911048

Query Match 6.8%; Score 507.1; DB 1; Length 628;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 565; Conservative 9; Mismatches 48; Indels 9; Gaps 3;

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||||
Db        628 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 569

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        568 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 509

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        508 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 449

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        448 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 390

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||   ||  | || ||||||||||||||||||||||||
Db        389 TCTACACTGAACAAGATTTATACA---ATCACGTACCTAAGCCCCGCAACAAAAGAGTAC 333

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||| : ||||||||||||||||||||
Db        332 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCRGACTAGGTACTGGCATTGGCA 273

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||||| |||||||
Db        272 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 213

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        212 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 153

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        152 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 93

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         92 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 33

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         32 AAGAAATTTGAGATTGATTACAACATAGAAC 2
```

RESULT 175
US-10-301-480-911050/c
; Sequence 911050, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911050
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-911050

Query Match 6.8%; Score 507.1; DB 1; Length 628;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 565; Conservative 9; Mismatches 48; Indels 9; Gaps 3;

```
Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||  | ||||||||||:||| | ||||||| || |||||| ||||||||||||||||||
Db        628 GGACTCCCTGTTTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTT 569

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||  | | |||| |||||||||||| |||||||||| || |||||||||| |
Db        568 AGCAATACTTTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGA 509

Qy       6361 ATAGTCTGCCTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTT 6415
             |||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db        508 ATAGTCTGCCTACCCTCAGGAATACTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTT 449

Qy       6416 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCA 6475
             |||||||||||||||||||||||||||||||||||||||||||||| |||| |||:|||
Db        448 GTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCA 390

Qy       6476 TCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTAC 6535
             |||||||||||||||||||||||:| |   | || ||||||||||||||||||||||||
Db        389 TCTACACTGAACAAGATTTATACARTCA---CGTACCTAAGCCCCGCAACAAAAGAGTAC 333

Qy       6536 CCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCG 6595
             |||||||||||||||||| |||||||||||||||||||||    ||||||||||||||||||
Db        332 CCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCA 273

Qy       6596 GTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGG 6655
             |||||||||||||| |||||||||||||||||||||||||||||||||||||| |||||||
Db        272 GTATCACAACCTCTGCTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGG 213

Qy       6656 AACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAG 6715
             || ||||  | |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        212 AATGGGTTACTGACTCCCTGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAG 153

Qy       6716 TCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTAT 6775
             ||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||||||||
Db        152 TCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTAT 93

Qy       6776 TTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTR 6835
             | |||||| |||||||||||||||||||| |||||||| |||||| | ||||||||||||:
Db         92 TATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTA 33

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGC 6866
             |||||||| :|||| || |||||| ||:| |
Db         32 AAGAAATTTGAGATTGATTACAACATAGAAC 2
```

```
RESULT 176
US-10-632-793-18
; Sequence 18, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
;  APPLICANT: PARANHOS-BACCALA, Glaucia
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  VOISSET, Cecile
;  TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
;  TITLE OF INVENTION:  AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
;  FILE REFERENCE: 110048
;  CURRENT APPLICATION NUMBER: US/10/632,793
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US/09/869,927
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: PCT/FR00/00144
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: FR 99/00888
;  PRIOR FILING DATE: 1999-01-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 678
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (594)..(594)
;   OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-18

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| |||||||||||||||:||:|||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:||||||||||||:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||              ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:||||||||||||||||:||||||||||||||:::||||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||||| |||||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:|||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:||||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             ||||||||||||||:|||:|||||: ||||||||||::||||::|||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
              |:||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 177
US-10-717-580-29
; Sequence 29, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEA
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
;   LENGTH: 678
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Ppol-MSRV probe
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (594)..(594)
;   OTHER INFORMATION: n = any nucleotide
US-10-717-580-29

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| |||||||||||||||:||:|||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:||||||||||||:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||              ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:||||||||||||||||:||||||||||||||:::||||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
```

```
                ||||||||| |||:|||||||||||| ||||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
                |||:||||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
                ||||:::| |  |||||||||||    ||  |   |||  ||||||||:||||||||||
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
                |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
                |||||||||||||||:|||:|||||: |||||||||||::||||::||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
                |:||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 178
US-09-925-065A-770753
; Sequence 770753, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 770753
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-770753

  Query Match             6.7%;  Score 502.7;  DB 1;  Length 628;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  566;  Conservative    4;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
                ||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db          1 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 59

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
                | |||||||||||| |||||||||||||||||||||||||||||||||||||| | | ||
Db         60 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 118

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
                || ||||||||||||||| |||||||||||||||||||||  ||||||||||||||||
Db        119 TCCCTGTCACCTGACTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 178

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
                ||||||||||||||||| ||||||||||||||||||  || ||| || ||||||||||||
Db        179 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCWCCTTAGTCCTGGAGCAAAACTTA 238

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
                |||||||||||||| ||||||||:| ||||||||||||||||||||| |||||||||| |
Db        239 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 298

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
                |||   ||||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        299 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 358

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
                |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        359 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 418

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
                |||||||| ||||||||||||||||||||||||||||||||| ||| ||||  |||
Db        419 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 478

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
                || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        479 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 538

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
                ||||||  |||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        539 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 598

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCCT 2079
                ||||||||| ||| ||||||||| ||||||
Db        599 GGGGCAAGCACCAGCCCATGCCAACACCCT 628


RESULT 179
US-09-925-065A-770753
; Sequence 770753, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 770753
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-770753

  Query Match             6.7%;  Score 502.7;  DB 1;  Length 628;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  566;  Conservative    4;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
                ||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db          1 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 59

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
                | |||||||||||| |||||||||||||||||||||||||||||||||||||| | | ||
Db         60 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 118
```

```
Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || |||||||||||||||| |||||||||||||||||||||||   ||||||||||||||||
Db        119 TCCCTGTCACCTGACTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 178

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             |||||||||||||||||  |||||||||||||||||   || ||| || ||||||||||||||
Db        179 TCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCCWCCTTAGTCCTGGAGCAAAACTTA 238

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| |||||||||||||||||||||| |||||||||| |
Db        239 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 298

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   ||||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        299 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 358

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        359 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 418

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             ||||||||| |||||||||||||||||||||||||||||||||||||| ||| |||| |||
Db        419 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 478

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        479 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 538

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        539 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 598

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCCT 2079
             ||||||||| ||| ||||||||| ||||||
Db        599 GGGGCAAGCACCAGCCCATGCCAACACCCT 628


RESULT 180
US-09-925-065A-835279
; Sequence 835279, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 835279
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-835279

  Query Match             6.7%;  Score 502.3;  DB 1;  Length 628;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  565;  Conservative    5;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             ||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db          1 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 59

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||  |||||||||||||||||||||||||||||||||||||||| | | ||
Db         60 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 118

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || ||||||||||:||| |||||||||||||||||||||||  ||||||||||||||||
Db        119 TCCCTGTCACCTGAYTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 178

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             |||||||||||||||||  |||||||||||||||||   || ||| || ||||||||||||||
Db        179 TCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCCACCTTAGTCCTGGAGCAAAACTTA 238

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| |||||||||||||||||||||| |||||||||| |
Db        239 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 298

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   ||||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        299 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 358

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        359 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 418

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             ||||||||| |||||||||||||||||||||||||||||||||||||| ||| |||| |||
Db        419 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 478

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        479 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 538

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        539 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 598

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCCT 2079
             ||||||||| ||| ||||||||| ||||||
Db        599 GGGGCAAGCACCAGCCCATGCCAACACCCT 628


RESULT 181
US-09-925-065A-835279
; Sequence 835279, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 835279
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-835279

  Query Match             6.7%;  Score 502.3;  DB 1;  Length 628;
```

```
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  565;  Conservative    5;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             ||||||||||:|||| ||: |||:|||||| ||| |||||||||  ||||||||||| |
Db             1 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 59

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||| ||||||||||||||||||||||||||||||||||||||| | | ||
Db            60 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 118

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || ||||||||||:||| |||||||||||||||||||||  ||||||||||||||||
Db        119 TCCCTGTCACCTGAYTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 178

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             |||||||||||||||| |||||||||||||||  || ||| || ||||||||||||||
Db        179 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCACCTTAGTCCTGGAGCAAAACTTA 238

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| |||||||||||||||||||||| |||||||||| |
Db        239 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 298

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   |||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        299 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 358

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        359 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 418

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             |||||||| |||||||||||||||||||||||||||||||||| ||| ||||  |||
Db        419 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 478

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        479 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 538

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        539 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 598

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCCT 2079
             ||||||||| ||| ||||||||| ||||||
Db        599 GGGGCAAGCACCAGCCCATGCCAACACCCT 628


RESULT 182
US-09-925-065A-760795/c
; Sequence 760795, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 760795
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-760795

  Query Match             6.7%;  Score 501.7;  DB 1;  Length 627;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  565;  Conservative    4;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             ||||||||||:|||| ||: |||:|||||| ||| |||||||||  ||||||||||| |
Db         627 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 569

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||| ||||||||||||||||||||||||||||||||||||||| | | ||
Db         568 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 510

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || |||||||||||||| |||||||||||||||||||||  ||||||||||||||||
Db         509 TCCCTGTCACCTGACTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 450

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             |||||||||||||||| |||||||||||||||  || ||| || ||||||||||||||
Db         449 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCWCCTTAGTCCTGGAGCAAAACTTA 390

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| |||||||||:| |||||||||||||||||||||| |||||||||| |
Db         389 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 330

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   |||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db         329 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 270

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         269 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 210

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             |||||||| |||||||||||||||||||||||||||||||||| ||| ||||  |||
Db         209 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 150

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db         149 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 90

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          89 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 30

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCC 2078
             ||||||||| ||| ||||||||| |||||
Db          29 GGGGCAAGCACCAGCCCATGCCAACACCC 1


RESULT 183
US-09-925-065A-760795/c
; Sequence 760795, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 760795
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-760795

  Query Match             6.7%;  Score 501.7;  DB 1;  Length 627;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  565;  Conservative    4;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             ||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db        627 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 569

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||| ||||||||||||||||||||||||||||||||||||| | | ||
Db        568 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 510

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || |||||||||||||| ||||||||||||||||||||  ||||||||||||||||
Db        509 TCCCTGTCACCTGACTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 450

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             ||||||||||||||||| ||||||||||||||||  || ||| || ||||||||||||||
Db        449 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCWCCTTAGTCCTGGAGCAAAACTTA 390

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| |||||||||||||||||||| |||||||||| |
Db        389 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 330

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   ||||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        329 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 270

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        269 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 210

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             ||||||||| |||||||||||||||||||||||||||||||| ||| ||||  |||
Db        209 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 150

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        149 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 90

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||| ||||||||||||||||||||||||||||||||
Db         89 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 30

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCC 2078
             ||||||||| ||| ||||||||| |||||
Db         29 GGGGCAAGCACCAGCCCATGCCAACACCC 1


RESULT 184
US-09-925-065A-829558/c
; Sequence 829558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 829558
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-829558

  Query Match             6.7%;  Score 501.3;  DB 1;  Length 627;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  564;  Conservative    5;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             ||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db        627 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 569

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||| ||||||||||||||||||||||||||||||||||||| | | ||
Db        568 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 510

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || |||||||||||:||| |||||||||||||||||||||  ||||||||||||||||
Db        509 TCCCTGTCACCTGAYTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 450

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             ||||||||||||||||| |||||||||||||||||  || ||| || ||||||||||||||
Db        449 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCACCTTAGTCCTGGAGCAAAACTTA 390

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| |||||||||||||||||||| |||||||||| |
Db        389 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 330

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   ||||||| |||| |||||||| ||||| |||||| ||||| |||||||| |||
Db        329 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 270

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        269 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 210

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             ||||||||| |||||||||||||||||||||||||||||||| ||| ||||  |||
Db        209 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 150

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||| |||| |  || ||||||
Db        149 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 90

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||| ||||||||||||||||||||||||||||||||
Db         89 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 30

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCC 2078
             ||||||||| ||| ||||||||| |||||
Db         29 GGGGCAAGCACCAGCCCATGCCAACACCC 1


RESULT 185
US-09-925-065A-829558/c
; Sequence 829558, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 829558
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-829558

  Query Match             6.7%;  Score 501.3;  DB 1;  Length 627;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  564;  Conservative    5;  Mismatches   57;  Indels    3;  Gaps    3;

Qy       1451 AAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGA 1510
             |||||||||||:|||| ||: |||:|||||| ||| |||||||||  |||||||||| |
Db        627 AAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAGTGATGTCCAGTATAACACAGGGAA 569

Qy       1511 AGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCC 1570
             | |||||||||||| |||||||||||||||||||||||||||||||||||||| | | ||
Db        568 A-GGAAGAAAATCTTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAACATACC 510

Qy       1571 TCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAG 1630
             || ||||||||||||:||| |||||||||||||||||||||||  ||||||||||||||||
Db        509 TCCCTGTCACCTGAYTCTACTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAG 450

Qy       1631 TCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTA 1689
             |||||||||||||||| |||||||||||||||  || ||| || ||||||||||||||
Db        449 TCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCACCTTAGTCCTGGAGCAAAACTTA 390

Qy       1690 GAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCG 1749
             |||||||||||||| ||||||||:| ||||||||||||||||||||| |||||||||| |
Db        389 GAAACCCTATTGAATTTGGCAACCTTGGTTTTTTATAATAGAGATCAAGAGGAGCAGGTG 330

Qy       1750 GAACAGGACAAACGGGATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
             |||   ||||||| |||| |||||||| ||||| |||||| ||||| ||||||||  |||
Db        329 GAATGAGACAAACAGGATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGC 270

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        269 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTG 210

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
             |||||||| |||||||||||||||||||||||||||||||||||||||| ||| ||||  |||
Db        209 CTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCC 150

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || || |||||||||||||| |||||||||||||||||||| |||| |  || ||||||
Db        149 CCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATG 90

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||||||   |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db          89 AAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 30

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCC 2078
             ||||||||| ||| |||||||||  |||||
Db          29 GGGGCAAGCACCAGCCCATGCCAACACCC 1


RESULT 186
US-09-925-065A-944021
; Sequence 944021, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944021
;   LENGTH: 607
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944021

  Query Match             6.7%;  Score 500.1;  DB 1;  Length 607;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  523;  Conservative   30;  Mismatches   51;  Indels    3;  Gaps    1;

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:| ||| |||||||| ||||||||:||||||||:|||||||||||||||
Db          1 TAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAATATATATACAG 60

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:| ||||:|||||||| ||||||||||||::|||||||:|:||||||||||||
Db         61 ACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAAT 120

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             |||||||||||:|||:|||||||||||||:|||||||||||||||||||:|||||||:|||
Db         121 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGG 180

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:|||||||:||| |||||||||:|||||||| |||||||||:
Db         181 CTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAGAAAGGAAAGG 240

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|| :  |||||  :|||||||| :|||||||  |||||| ||
Db         241 AAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCCACAAGGCTGG 300

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||||    |   ||| |||||| |||||||||| ||:|||||
Db         301 ACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCCCTCYGGGAAG 360

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||||||| ||||| |||||| ||| ||||    | ||  |
Db         361 CCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAGGGAACTTCATGAGGACATACTTCCC 420

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAAT 4508
             |||||||  || || |||||||  |  |||| |||||||||||||||||| ||| |||||
Db         421 TCCCCTCCAGATGGCTAGCCACCAATAAAGGAAAAATACTTTTGCCTGCAGCTAACCAAT 480

Qy       4509 GGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARA 4568
              |||||||||||||||||||||||||||||||| |||||||||| |||||||||||  |:|
Db         481 AGAAATTACTTAAAACCCTTCATCAAACCTTCCACTTAGGCATTGATAGCACCCATGAGA 540

Qy       4569 TGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGG 4628
             ||||||||| |||||||||||||||||||||||||||||||||||||||:|||:||||||
Db         541 TGGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGG 600

Qy       4629 CCTGTGA 4635
             ||||| |
Db         601 CCTGTAA 607
```

```
RESULT 187
US-09-925-065A-944021
; Sequence 944021, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944021
;   LENGTH: 607
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944021

  Query Match             6.7%;  Score 500.1;  DB 1;  Length 607;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  523;  Conservative   30;  Mismatches   51;  Indels    3;  Gaps    1;

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:| ||| |||||||| |||||||:|||||||:|||||||||||||||
Db          1 TAATAGCCCTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGGGTAAATATATATACAG 60

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:| ||||:|||||||| ||||||||||||::|||||||:|:||||||||||
Db         61 ACTCTAAGTGTGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAAT 120

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:||||||||||||:|||||||||||||||||||:|||||||:|||
Db        121 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAAATTATTATTGG 180

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             |:|||||||||||||:|||||||:||| |||||||||:|||||||| |||||||||:
Db        181 CTGTACAGAAACCTAAAGAGGTGGCAGTTTTACACTGCCGGGGTCATCAGAAAGGAAAGG 240

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|| :  ||||| :|||||||| :|||||||  |||||| ||
Db        241 AAAGGGAAATACAAGGGAGCCACCAAGTTGATATTGAAGTCAAAAGAGCCACAAGGCTGG 300

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||    |   ||| |||||| |||||||||| ||:|||||
Db        301 ACCCTCCATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAATCCCCTCYGGGAAG 360

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
             |||||||||||||||||||||||||||| ||||| |||||| ||| ||||    | ||  |
Db        361 CCAAGCCCCAGTACTCAGCAGGAGAAATAGAATAGGGAACTTCATGAGGACATACTTCCC 420

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAAT 4508
             |||||||  || || |||||||  |  |||| |||||||||||||||||| ||| |||||
Db        421 TCCCCTCCAGATGGCTAGCCACCAATAAAGGAAAAATACTTTTGCCTGCAGCTAACCAAT 480

Qy       4509 GGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARA 4568
              |||||||||||||||||||||||||||||| |||||||||| |||||||||||  |:|
Db        481 AGAAATTACTTAAAACCCTTCATCAAACCTTCCACTTAGGCATTGATAGCACCCATGAGA 540

Qy       4569 TGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGG 4628
             |||||||| |||||||||||||||||||||||||||||||||||||||:|||:||||||
Db        541 TGGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAGGG 600

Qy       4629 CCTGTGA 4635
             ||||| |
Db         601 CCTGTAA 607


RESULT 188
US-09-925-065A-933320/c
; Sequence 933320, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 933320
;   LENGTH: 620
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-933320

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 620;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  523;  Conservative   42;  Mismatches   49;  Indels    2;  Gaps    2;

Qy       3398 AACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
             |||  |||:|||||||||||:|| |||||:|||||||||||||:||:|||::|||||||
Db        615 AACTTAAGCCCCAGTGTTAAGCTTCCCAACAGGGCAAGACTTTTCTTTATATGTCACAGA 556

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
             |||||||| :||:||||| :|||||||||||| ||:|| :|| ||:|||||| ||||::
Db        555 AAAAAACAAGAATAGCTCCAGGAGTCCTTACATAGGTCTGAGAGATGAGCTTGGAACCCA 496

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
             ||||::|||||:||||||||:||||| |||||||||||||| |:||||||||||:||
Db        495 TGGCATACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGTTGTCCTCATTGTTTATGGA 436

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
             ||||||  |||||| || ||::  ||||||  |||||||||||||||||| ||:|| ||||
Db        435 TAGTGGCAGCAGTAACAATCTAGGTATCTTTAGCAGTTAAAATAATACAAGGAAGGGATC 376

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
             ||||||||||||||||||||:||||||::|||||||||||||||||||||||||||||||
Db        375 TTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 316

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
             ||||||||||: ||||||||||:|||||||||||| |||||: || ||||||| :|
Db        315 TGTCAGACAACTTTTTACTTAAATATCAGGCTCTATTATTTGAAGTGCTAGTGCTGTGAA 256

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
             ||:|||||||||||||||||||||:  |||||||||||||| ||||||||||:
Db        255 TGYGCACTTGTGCAACTCTTAACCCAGCTACATTTCTTCCAGACAATAAAGAAAAGATAG 196

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
             ||:|||||||||||:| ||| |||||||  ||  ||||| |||||||  ||:||||:|
Db        195 AACATAACTGTCAACAAGTGATTGCTCAAACATACACCACTAGAGGGGAGGTTTTAGAGG 136

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
              |||:|| ||||||||: ||| ||||||||||||||||||||||||||||||||||||
Db        135 ATCCCTTCACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 77
```

```
Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA-GTAATCC 3996
Db         76 CGACTTTGAATAGCAGGGTATGTGGTGGTCAGTGATAATGGAATACTTGAAAGGTAATCC 17

Qy       3997 CCTCACTCCAGGAACT 4012
Db         16 CCTAACTCCAGGAACT 1
```

RESULT 189
US-09-925-065A-933320/c
; Sequence 933320, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 933320
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-933320

Query Match 6.7%; Score 498.4; DB 1; Length 620;
Best Local Similarity 84.9%; Pred. No. 0;
Matches 523; Conservative 42; Mismatches 49; Indels 2; Gaps 2;

```
Qy       3398 AACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
Db        615 AACTTAAGCCCCAGTGTTAAGCTTCCCAACAGGGCAAGACTTTTCTTTATATGTCACAGA 556

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
Db        555 AAAAAACAAGAATAGCTCCAGGAGTCCTTACATAGGTCTGAGAGATGAGCTTGGAACCCA 496

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
Db        495 TGGCATACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGTTGTCCTCATTGTTTATGGA 436

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
Db        435 TAGTGGCAGCAGTAACAATCTAGGTATCTTTAGCAGTTAAAATAATACAAGGAAGGGATC 376

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
Db        375 TTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 316

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
Db        315 TGTCAGACAACTTTTTACTTAAATATCAGGCTCTATTATTTGAAGTGCTAGTGCTGTGAA 256

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
Db        255 TGYGCACTTGTGCAACTCTTAACCCAGCTACATTTCTTCCAGACAATAAAGAAAAGATAG 196

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
```

```
Db        195 AACATAACTGTCAACAAGTGATTGCTCAAACATACACCACTAGAGGGGAGGTTTTAGAGG 136

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
Db        135 ATCCCTTCACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 77

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA-GTAATCC 3996
Db         76 CGACTTTGAATAGCAGGGTATGTGGTGGTCAGTGATAATGGAATACTTGAAAGGTAATCC 17

Qy       3997 CCTCACTCCAGGAACT 4012
Db         16 CCTAACTCCAGGAACT 1
```

RESULT 190
US-09-925-065A-923906
; Sequence 923906, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923906
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-923906

Query Match 6.6%; Score 496; DB 1; Length 608;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 520; Conservative 50; Mismatches 36; Indels 6; Gaps 3;

```
Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
Db          1 ATACTGGCTTATCCTCACCCCAAAACCCTAAAGCAACTAAGACGGTTCCTTGGCATAACA 60

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
Db         61 GGTTTCTGCCGAATATGGATT-TCCAGGTACAGTGAAATAGCCAGACCATTATATACACT 119

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
Db        120 AATTAAGGAAACTCAGAAAGCCAATATCCATTTAGTAAGATGGATACCTGAAACAGAAGC 179

Qy       3369 GGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACR 3428
Db        180 AGCTTTCCAGGCCCTAAAGAAGGCCCT---AACCAAGCCCCAGTGTTAAGCTTGCCAATG 236

Qy       3429 GGGCAAGACTTTTSTTYATAYRT--CACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
Db        237 GGGCAAGACTTTTYTTTATATGTCACACACAAAAAAACAGGAATAGCTCTAAGAGTCCTT 296

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
Db        297 ACACAGGTCCAAGAGATGAGTTTGCAACCCGTGGCATACCTGAGTAAGGAAACTGATGCA 356

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
```

```
              |||||||||||||||:|:||||||||||: |||||||  |||||||  ||||::||||||||
Db        357 GTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGACGGCAGTAACAGTCTTAGTATCT 416

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              ||||||||||||||||||||||:|||||| ||||||||||||||||||||:|||||:
Db        417 GAAGCAGTTAAAATAATACAGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAAT 476

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :|||||||||||||||||||||||||||||| |||||||||  ||||||||||||:||||
Db        477 GGCATACTCACTGCTAAAGGAGACTTGTGGCCGTCAGACAAAAGTTTACTTAAATATCAG 536

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |||||||||||||||:||||||||||||:|||| |||||||||||||||||||||||| |
Db        537 GCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGGC 596

Qy       3787 NCATTTCTTCCA 3798
              |||||||| ||
Db        597 ACATTTCTTTCA 608


RESULT 191
US-09-925-065A-923906
; Sequence 923906, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923906
;   LENGTH: 608
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-923906

  Query Match             6.6%;  Score 496;  DB 1;  Length 608;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  520;  Conservative   50;  Mismatches   36;  Indels    6;  Gaps    3;

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||||| |:|||||||||||||||||||: ::|||||||:|||:|
Db          1 ATACTGGCTTATCCTCACCCCAAAACCCTAAAGCAACTAAGACGGTTCCTTGGCATAACA 60

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||||||:|||||||  ||||||:  | :|||||||||| :|||||:|||||:|
Db         61 GGTTTCTGCCGAATATGGATT-TCCAGGTACAGTGAAATAGCCAGACCATTATATACACT 119

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||| |||||||:|||||||:|:|||||  :::|||
Db        120 AATTAAGGAAACTCAGAAAGCCAATATCCATTTAGTAAGATGGATACCTGAAACAGAAGC 179

Qy       3369 GGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACR 3428
              ||||||||||||  ||  | | |||   | |||||:||||||||||:|||||||  :
Db        180 AGCTTTCCAGGCCCTAAAGAAGGCCCT---AACCAAGCCCCAGTGTTAAGCTTGCCAATG 236

Qy       3429 GGGCAAGACTTTTSTTYATAYRT--CACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
              ||||||||||||:||:|||::|  |||| ||||||||||:||:||||||: ||||||||
Db        237 GGGCAAGACTTTTYTTTATATGTCACACACAAAAAAACAGGAATAGCTCTAAGAGTCCTT 296
```

```
Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
              ||||||:|||:|| ||:||| |||||||||::||||::|||||||:||||||||:||||| |
Db        297 ACACAGGTCCAAGAGATGAGTTTGCAACCCGTGGCATACCTGAGTAAGGAAACTGATGCA 356

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
              |||||||||||||:|:||||||||||: |||||||  |||||||  ||||::||||||||
Db        357 GTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGACGGCAGTAACAGTCTTAGTATCT 416

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              ||||||||||||||||||||||||:|||||| |||||||||||||||||||||||:|||||:
Db        417 GAAGCAGTTAAAATAATACAGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAAT 476

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :|||||||||||||||||||||||||||||| |||||||||  ||||||||||||:||||
Db        477 GGCATACTCACTGCTAAAGGAGACTTGTGGCCGTCAGACAAAAGTTTACTTAAATATCAG 536

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |||||||||||||||:||||||||||||:|||| |||||||||||||||||||||||| |
Db        537 GCTCTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGGC 596

Qy       3787 NCATTTCTTCCA 3798
              |||||||| ||
Db        597 ACATTTCTTTCA 608


RESULT 192
US-10-301-480-441252/c
; Sequence 441252, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 441252
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-441252

  Query Match             6.6%;  Score 495.1;  DB 1;  Length 618;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  525;  Conservative   29;  Mismatches   61;  Indels    3;  Gaps    1;

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              | |||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db         618 AATTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 559

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||| |||||||||||||||||  || |||:||||:| ||||| ||||||
Db         558 ATACTTGAAAATAATCCCCTCACTCCAGGAAGAAGCGCTTAGCTGGAAGAACGAATAGCT 499

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRAR 4099
              :|||:|  |||| |||||||||||||:||||||||:|||||||||||||||||||:|:
Db         498 CTCACTCAGGCAATAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAG 439

Qy       4100 TATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACT 4159
              ||||||:| |||||| ||||||||||  ::|||||||:|:|||||| ||||||| |||
Db         438 TATGCTTAGCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGAAATTCCTGACT 379

Qy       4160 TCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAG 4219
              ||: | :|||||||  ||||:|||||||||||||||  ||||:||||||| |||:||||||
Db         378 TCCAAAGGAACACCCATCAAACATCAGGAAGCCATCAGGAGATTATTAYCGGCTGTACAG 319

Qy       4220 AAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAA 4279
```

```
                |||||||:||||||||:| |||||  ||||:||||||||||| ||||||||  :||||||||
Db        318 AAACCTAAAGAGGTGGCAATCTTATGCTGCCGGGGTCATCAGAAAGGAAAAGAAAGGGAA 259

Qy       4280 ATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCA 4339
                |||:|||:|||:|||||| | :||||  ||||:|||| |||||||||| ||||||||||
Db        258 ATAGAAGGGAACTGCCAAACGGATATCAAAGCCAAAATAGCTGCAAGGTAGGACCCTCCA 199

Qy       4340 TTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCC 4399
                |||||||| |||||  ||   ||| |||||| |||||||||| || ||||||||||||||
Db        198 TTAGAAATACTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCC 139

Qy       4400 CAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCG 4456
                |||||||||||||||||| |||||||||| |||| ||||    ||||| ||| ||||
Db        138 CAGTACTCAGCAGGAGAAATAGAATGGGGAAACTCATGAGGACATAGTTTCCTCTCCTCA 79

Qy       4457 GGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTA 4516
                ||| || ||| |||  |||||| |||||||||||||||| ||| ||||||||||| |
Db         78 GGATGGCTAGGCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATGA 19

Qy       4517 CTTAAAACCCTTCATCAA 4534
                ||||||||||||| |||
Db         18 CTTAAAACCCTTCACCAA 1


RESULT 193
US-10-301-480-1054661/c
; Sequence 1054661, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1054661
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1054661

  Query Match             6.6%;  Score 495.1;  DB 1;  Length 618;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  525;  Conservative   29;  Mismatches   61;  Indels    3;  Gaps    1;

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
                | |||||||||||||||||||| ||||||:|||||||||||||||||||||||||||||
Db         618 AATTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 559

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
                |||:|||||| ||||||||||||||||||||  || |||:||||:| ||||| ||||||
Db         558 ATACTTGAAAATAATCCCCTCACTCCAGGAAGAAGCGCTTAGCTGGAAGAACGAATAGCT 499

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRAR 4099
                :|||:|  |||| |||||||||||||:||||||||:|||||||||||||||||||:|:
Db         498 CTCACTCAGGCAATAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAG 439

Qy       4100 TATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACT 4159
                ||||||:| |||||| ||||||||||  ::|||||||:|:|||||| |||||| |||
Db         438 TATGCTTAGCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGAAATTCCTGACT 379

Qy       4160 TCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAG 4219
                ||: | :||||||| ||||:|||||||||||||| ||||:|||||||| |||:||||||
Db         378 TCCAAAGGAACACCCATCAAACATCAGGAAGCCATCAGGAGATTATTAYCGGCTGTACAG 319

Qy       4220 AAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAA 4279
                |||||||:|||||||||:| |||||  ||||:|||||||||| |||||||| :||||||||
```

```
Db        318 AAACCTAAAGAGGTGGCAATCTTATGCTGCCGGGGTCATCAGAAAGGAAAAGAAAGGGAA 259

Qy       4280 ATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCA 4339
                |||:|||:|||:|||||| | :||||  ||||:|||| |||||||||| ||||||||||
Db        258 ATAGAAGGGAACTGCCAAACGGATATCAAAGCCAAAATAGCTGCAAGGTAGGACCCTCCA 199

Qy       4340 TTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCC 4399
                |||||||| |||||  ||   ||| |||||| |||||||||| || ||||||||||||||
Db        198 TTAGAAATACTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCC 139

Qy       4400 CAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCG 4456
                |||||||||||||||||| |||||||||| |||| ||||    ||||| ||| ||||
Db        138 CAGTACTCAGCAGGAGAAATAGAATGGGGAAACTCATGAGGACATAGTTTCCTCTCCTCA 79

Qy       4457 GGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTA 4516
                ||| || ||| |||  |||||| |||||||||||||||| ||| ||||||||||| |
Db         78 GGATGGCTAGGCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATGA 19

Qy       4517 CTTAAAACCCTTCATCAA 4534
                ||||||||||||| |||
Db         18 CTTAAAACCCTTCACCAA 1


RESULT 194
US-10-301-480-475703
; Sequence 475703, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 475703
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-475703

  Query Match             6.6%;  Score 494.6;  DB 1;  Length 609;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  527;  Conservative   29;  Mismatches   45;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
                |||||||||||||||||||||||||||||||||| |||||:|||||||||||||||||||
Db           1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
                ||||||||||:|||||||||||||||||||||| ||||||||||:||||:||||||||
Db          61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
                |||||||:|||:| |||||||||||||||||||:|||||||:|||||||||||||||||
Db         121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
                | || |: |       |||: |||||||| ||||||||||||::|||||||:|:|||||||
Db         181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGASAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
                |||||||||||||:|||: ||||||||||:|||||||||||||||| ||:||||| |
Db         241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCA 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
                :||||:|||||||||||:|||||||| :||||||||||||: |||||||| |||| |
Db         301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360
```

```
Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||| |:|| :|||:  |||||| :||||||||:|||||||||||||||||
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||| |||||||||  ||   ||| |||||| ||||||| || ||| |
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             ||||||||||||||||||||||||||||| ||||||||||||||| |||     |||
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
             || |||||||| ||| || ||||||| |||||||||||||||||| ||||||| ||| |
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGCAGCTAGC 600

Qy       4505 CAATGGAAA 4513
             ||| |||||
Db        601 CAACGGAAA 609


RESULT 195
US-10-301-480-1089112
; Sequence 1089112, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1089112
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1089112

  Query Match             6.6%;  Score 494.6;  DB 1;  Length 609;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  527;  Conservative   29;  Mismatches   45;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||||||||||||||||||||| |||||||:|||||||||||||||||||||
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||||||||||||| ||||||||||||:||||:||||||||
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| ||||||||||||||||||||:|||||||||:|||||||||||||||
Db        121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |       |||: |||||||| ||||||||||::||||||||:|:|||||||
Db        181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGASAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             ||||||||||||||:|||: ||||||||||:|||||||||||||||||| ||:||||| |
Db        241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCA 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:|||||||||||:||||||| :|||||||||||||: |||||||| |||| |
Db        301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||| |:|| :|||:  |||||| :||||||||:|||||||||||||||||
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||| |||||||||  ||   ||| |||||| ||||||| || ||| |
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             ||||||||||||||||||||||||||||| ||||||||||||||| |||     |||
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
             || |||||||| ||| || ||||||| |||||||||||||||||| ||||||| ||| |
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGCAGCTAGC 600

Qy       4505 CAATGGAAA 4513
             ||| |||||
Db        601 CAACGGAAA 609


RESULT 196
US-10-301-480-475704
; Sequence 475704, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 475704
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-475704

  Query Match             6.6%;  Score 494.2;  DB 1;  Length 609;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  526;  Conservative   30;  Mismatches   45;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||||||||||||||||||||| |||||||:|||||||||||||||||||||
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||||||||||||| ||||||||||||:||||:||||||||
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| ||||||||||||||||||||:|||||||||:|||||||||||||||
Db        121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |       |||: |||||||| ||||||||||::||||||||:|:|||||||
Db        181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             ||||||||||||||:|||: ||||||||||:|||||||||||||||||| ||:||||| :
Db        241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCR 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:|||||||||||:||||||| :|||||||||||||: |||||||| |||| |
Db        301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
```

```
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             ||||||||||||||| |||||||||| ||   ||| |||||| ||||||| || ||| |
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             |||||||||||||||||||||||||||||| |||||||||||||||| |||      |||
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
             || |||||||| ||| || ||||||| |||||||||||||||||| ||||||| ||| |
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGCAGCTAGC 600

Qy       4505 CAATGGAAA 4513
             ||| |||||
Db        601 CAACGGAAA 609


RESULT 197
US-10-301-480-1089113
; Sequence 1089113, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1089113
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1089113

  Query Match             6.6%;  Score 494.2;  DB 1;  Length 609;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  526;  Conservative   30;  Mismatches   45;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||||||||||||||||||||| |||||||:|||||||||||||||||||||
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             |||||||||:|||||||||||||||||||||||| |||||||||||:||||:||||||||
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| |||||||||||||||||||:||||||||:|||||||||||||||||
Db         121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |       |||: ||||||| ||||||||||||::||||||||:|:|||||||
Db         181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             ||||||||||||||:|||: |||||||||||:|||||||||||||||| ||:||||| :
Db         241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCR 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:||||||||||||:||||||| :||||||||||||: |||||||| |||| |
Db         301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||||| |:|| :|||:  |||||| :||||||||||:|||||||||||||||
```

```
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||||| |||||||||| ||   ||| |||||| ||||||| || ||| |
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             |||||||||||||||||||||||||||||| |||||||||||||||| |||      |||
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
             || |||||||| ||| || ||||||| |||||||||||||||||| ||||||| ||| |
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGCAGCTAGC 600

Qy       4505 CAATGGAAA 4513
             ||| |||||
Db        601 CAACGGAAA 609


RESULT 198
US-09-925-065A-371249/c
; Sequence 371249, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 371249
;   LENGTH: 613
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-371249

  Query Match             6.6%;  Score 491.7;  DB 1;  Length 613;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  521;  Conservative   29;  Mismatches   60;  Indels    3;  Gaps    1;

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYT 3984
             ||||||||||||||||||| |||||:|||||||||||||||||||||||||||||||||:|
Db         613 CTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACT 554

Qy       3985 TGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAY 4044
             ||||| |||||||||||||||||  || |||:||||:| ||||| |||||| :|||:
Db         553 TGAAAATAATCCCCTCACTCCAGGAAGAAGCGCTTAGCTGGAAGAACGAATAGCTCTCAC 494

Qy       4045 TKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGC 4104
             |  |||| ||||||||||||||:||||||||:|||||||||||||||||||:|:|||||
Db         493 TCAGGCAATAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGC 434

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
             |:| |||||| |||||||||  ::|||||||:|:|||||| |||||| |||||: |
Db         433 TTAGCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGAAATTCCTGACTTCCAA 374

Qy       4165 GRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACC 4224
              :|||||| ||||:|||||||||||||| ||||:|||||||| |||:||||||||||
Db         373 AGGAACACCCATCAAACATCAGGAAGCCATCAGGAGATTATTAYCGGCTGTACAGAAACC 314

Qy       4225 TARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASA 4284
```

```
              ||:|||||||||:| |||||  ||||:||||||||||  ||||||||  :||||||||||:|
Db        313 TAAAGAGGTGGCAATCTTATGCTGCCGGGGTCATCAGAAAGGAAAAGAAAGGGAAATAGA 254

Qy       4285 AGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGA 4344
              ||:|||:|||||| | :||||  ||||:|||| ||||||||||  ||||||||||||||||
Db        253 AGGGAACTGCCAAACGGATATCAAAGCCAAAATAGCTGCAAGGTAGGACCCTCCATTAGA 194

Qy       4345 AATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTA 4404
              ||| |||||  ||   ||| |||||| ||||||||| || ||||||||||||||||||||
Db        193 AATACTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTA 134

Qy       4405 CTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACG 4461
              |||||||||||||| |||||||||| |||| ||||    ||||| ||| |||| ||| |
Db        133 CTCAGCAGGAGAAATAGAATGGGGAAACTCATGAGGACATAGTTTCCTCTCCTCAGGATG 74

Qy       4462 GTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAA 4521
              | ||| |||  ||||||| |||||||||||||||||| ||| ||||||||||| ||||||
Db         73 GCTAGGCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATGACTTAA 14

Qy       4522 AACCCTTCATCAA 4534
              ||||||||| |||
Db         13 AACCCTTCACCAA 1


RESULT 199
US-09-925-065A-371249/c
; Sequence 371249, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 371249
;   LENGTH: 613
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-371249

  Query Match             6.6%;  Score 491.7;  DB 1;  Length 613;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  521;  Conservative   29;  Mismatches   60;  Indels    3;  Gaps    1;

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYT 3984
              |||||||||||||||||| |||||:|||||||||||||||||||||||||||||||:|
Db         613 CTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACT 554

Qy       3985 TGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAY 4044
              ||||| |||||||||||||||||  || |||:||||:| ||||| |||||| :|||:
Db         553 TGAAAATAATCCCCTCACTCCAGGAAGAAGCGCTTAGCTGGAAGAACGAATAGCTCTCAC 494

Qy       4045 TKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGC 4104
              |  |||| |||||||||||||:||||||||:|||||||||||||||||||:|:|||||
Db         493 TCAGGCAATAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGC 434

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
              |:| |||||| |||||||||  ::|||||||:|:||||||| ||||||| |||||: |
Db         433 TTAGCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGAAATTCCTGACTTCCAA 374

Qy       4165 GRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACC 4224
              :||||||| ||||:|||||||||||||| ||||:||||||| |||:||||||||||
Db         373 AGGAACACCCATCAAACATCAGGAAGCCATCAGGAGATTATTAYCGGCTGTACAGAAACC 314

Qy       4225 TARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASA 4284
              ||:|||||||:| |||||  |||||:|||||||||| |||||||| :|||||||||||:|
Db         313 TAAAGAGGTGGCAATCTTATGCTGCCGGGGTCATCAGAAAGGAAAAGAAAGGGAAATAGA 254

Qy       4285 AGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGA 4344
              ||:|||:|||||| | :||||  ||||:|||| ||||||||||  ||||||||||||||||
Db         253 AGGGAACTGCCAAACGGATATCAAAGCCAAAATAGCTGCAAGGTAGGACCCTCCATTAGA 194

Qy       4345 AATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTA 4404
              ||| |||||  ||   ||| |||||| ||||||||| || ||||||||||||||||||||
Db         193 AATACTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTA 134

Qy       4405 CTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACG 4461
              |||||||||||||| |||||||||| |||| ||||    ||||| ||| |||| ||| |
Db         133 CTCAGCAGGAGAAATAGAATGGGGAAACTCATGAGGACATAGTTTCCTCTCCTCAGGATG 74

Qy       4462 GTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAA 4521
              | ||| |||  ||||||| |||||||||||||||||| ||| ||||||||||| ||||||
Db         73 GCTAGGCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATGACTTAA 14

Qy       4522 AACCCTTCATCAA 4534
              ||||||||| |||
Db         13 AACCCTTCACCAA 1


RESULT 200
US-09-925-065A-946840
; Sequence 946840, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 946840
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-946840

  Query Match             6.5%;  Score 490.2;  DB 1;  Length 622;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  529;  Conservative   29;  Mismatches   54;  Indels   10;  Gaps    2;

Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
              |||||||||||||||:|||:| ||||||||||||||||||:|||||||:|||||||
Db           1 GCAGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATA 60

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
              ||||||| ||||:|:||||||:|||||||  ||||||||||:|::|||||||:|:||||
Db          61 TATACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAA 120

Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
              |||||||||||||||: ||:||||||||||||:|||||| |||||||||||| ||||
Db         121 AGGGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTA 180
```

```
Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT------ 4258
Db        181 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAA 240

Qy       4259 -CANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAG 4317
Db        241 AAAAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAG 300

Qy       4318 AGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAA 4377
Db        301 AGCCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAA 360

Qy       4378 TCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACG 4437
Db        361 CCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAG 420

Qy       4438 AGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCC 4494
Db        421 AGGACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCC 480

Qy       4495 TGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGA 4554
Db        481 TGCAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGA 540

Qy       4555 TAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 4614
Db        541 TAGCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 600

Qy       4615 GCARATAKTCAGGGCCTGTGAA 4636
Db        601 ATAGATAGTCAGGGCCTGTAAA 622
```

RESULT 201
US-09-925-065A-946840
; Sequence 946840, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 946840
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-946840

Query Match 6.5%; Score 490.2; DB 1; Length 622;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 529; Conservative 29; Mismatches 54; Indels 10; Gaps 2;

```
Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
Db          1 GCAGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATA 60

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
```

```
Db         61 TATACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAA 120

Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
Db        121 AGGGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTA 180

Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT------ 4258
Db        181 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAA 240

Qy       4259 -CANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAG 4317
Db        241 AAAAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAG 300

Qy       4318 AGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAA 4377
Db        301 AGCCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAA 360

Qy       4378 TCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACG 4437
Db        361 CCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAG 420

Qy       4438 AGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCC 4494
Db        421 AGGACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCC 480

Qy       4495 TGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGA 4554
Db        481 TGCAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGA 540

Qy       4555 TAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 4614
Db        541 TAGCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAA 600

Qy       4615 GCARATAKTCAGGGCCTGTGAA 4636
Db        601 ATAGATAGTCAGGGCCTGTAAA 622
```

RESULT 202
US-10-301-480-384873/c
; Sequence 384873, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 384873
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-384873

Query Match 6.5%; Score 490; DB 1; Length 618;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 534; Conservative 40; Mismatches 42; Indels 8; Gaps 4;

```
Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
Db        617 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 558

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
Db        557 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 498
```

```
Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        497 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 439

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||||| |||||:||||||| |||||||||||||||||||||||||:
Db        438 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 379

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:||||||||||||| :|||
Db        378 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 319

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             :|  || ||||||||||||||||: |||||||:||||||||||||| ||||:|:||||
Db        318 CTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 259

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||::||||||| | |||||||||::||||||||:|:||||||||||||||||||||||:
Db        258 CTTRCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 199

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:|||| |:|| |:||||||||||
Db        198 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 139

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:|||||||| :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        138 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 84

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||| :|||||||| ||||||||||| ||||||||
Db          83 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 25

Qy       4344 AAATGCTTATTAAACTTCCCTTAG 4367
             |||||||||| ||     | |||
Db          24 AAATGCTTATAGAAGGATGCCTAG 1
```

RESULT 203
US-10-301-480-998282/c
; Sequence 998282, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 998282
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-998282

Query Match 6.5%; Score 490; DB 1; Length 618;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 534; Conservative 40; Mismatches 42; Indels 8; Gaps 4;

```
Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             ||||||||| :|||| |||||||||| ||||||||||:| |||||||| ||||||||
Db        617 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 558

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| |||||||||||||| ||| ||||
Db        557 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 498
```

```
Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        497 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 439

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||||| |||||:||||||| |||||||||||||||||||||||||:
Db        438 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 379

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:||||||||||||| :|||
Db        378 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 319

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             :|  || ||||||||||||||||: |||||||:||||||||||||| ||||:|:||||
Db        318 CTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 259

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||::||||||| | |||||||||::||||||||:|:||||||||||||||||||||||:
Db        258 CTTRCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 199

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:|||| |:|| |:||||||||||
Db        198 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 139

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:|||||||| :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        138 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 84

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||| :|||||||| ||||||||||| ||||||||
Db          83 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 25

Qy       4344 AAATGCTTATTAAACTTCCCTTAG 4367
             |||||||||| ||     | |||
Db          24 AAATGCTTATAGAAGGATGCCTAG 1
```

RESULT 204
US-10-450-763-24195
; Sequence 24195, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 24195
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (3)..(173)
; OTHER INFORMATION: 92% homologous to Homo sapiens envelope protein,accession
; OTHER INFORMATION: number AF072508,Smith-Waterman Score=257.
US-10-450-763-24195

Query Match 6.5%; Score 489.8; DB 1; Length 583;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 515; Conservative 11; Mismatches 28; Indels 10; Gaps 1;

```
Qy       7019 TCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 7078
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCA 60
```

```
Qy       7079 GACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 7138
Db         61 GACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTC 120

Qy       7139 CTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGA 7198
Db        121 GTGAGGAAATCTCAGCTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGA 180

Qy       7199 GCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGA 7258
Db        181 GCGGTCGTCAGTCAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGA 240

Qy       7259 GAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGG 7318
Db        241 GAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGG 300

Qy       7319 TGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAG 7378
Db        301 TGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAG 360

Qy       7379 AGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATT 7438
Db        361 AGCTCACTAAAATGCTAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATC 420

Qy       7439 GCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGC 7498
Db        421 GCCTGAGAGCACAGCAAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGC 480

Qy       7499 AACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTC 7558
Db        481 AACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTC 530

Qy       7559 ACTCTATTAAATCTTGCARCTGCR 7582
Db        531 ACTCTATTAAATCTTGCAACTGCA 554
```

```
RESULT 205
US-09-925-065A-309261/c
; Sequence 309261, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 309261
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-309261

  Query Match             6.5%;  Score 489.6;  DB 1;  Length 614;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  530;  Conservative   40;  Mismatches   36;  Indels    8;  Gaps    4;

Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
Db        613 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 554

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
Db        553 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 494

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
Db        493 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 435

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
Db        434 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 375

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
Db        374 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 315

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
Db        314 CTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 255

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
Db        254 CTTRCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 195

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
Db        194 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 135

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
Db        134 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 80

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
Db         79 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 21

Qy       4344 AAATGCTTATTAAA 4357
Db         20 AAATGCTTATAGAA 7
```

```
RESULT 206
US-09-925-065A-309261/c
; Sequence 309261, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 309261
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-309261

  Query Match             6.5%;  Score 489.6;  DB 1;  Length 614;
```

```
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  530;  Conservative   40;  Mismatches   36;  Indels    8;  Gaps    4;

Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             ||||||||| :|||| |||||||||| |||||||||||:| |||||||| ||||||||
Db        613 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 554

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| ||||||||||||| ||| ||||
Db        553 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 494

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||
Db        493 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 435

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||| |||||:||||||| |||||||||||||||||||||||||:
Db        434 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 375

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:|||||||||||||| :|||
Db        374 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 315

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             :|  || ||||||||||||||||: |||||||:||||||||||||| ||||:|:||||
Db        314 CTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 255

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||::|||||||  | ||||||||||::|||||||:|:|||||||||||||||||||||:
Db        254 CTTRCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 195

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:||||  |:|| |:||||||||||
Db        194 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 135

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:|||||||  :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        134 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 80

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||  :|||||||  ||||||||||| ||||||||
Db          79 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 21

Qy       4344 AAATGCTTATTAAA 4357
             ||||||||||  ||
Db          20 AAATGCTTATAGAA 7


RESULT 207
US-10-301-480-384874/c
; Sequence 384874, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 384874
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-384874

  Query Match             6.5%;  Score 488.8;  DB 1;  Length 618;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  535;  Conservative   38;  Mismatches   43;  Indels    8;  Gaps    4;

Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             ||||||||| :|||| |||||||||| |||||||||||:| |||||||| ||||||||
Db        617 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 558

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| ||||||||||||| ||| ||||
Db        557 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 498

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||
Db        497 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 439

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||| |||||:||||||| |||||||||||||||||||||||||:
Db        438 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 379

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:|||||||||||||| :|||
Db        378 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 319

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             ||  || ||||||||||||||||: |||||||:||||||||||||| ||||:|:||||
Db        318 YTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 259

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||: ||||||| | ||||||||||::|||||||:|:|||||||||||||||||||||:
Db        258 CTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 199

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:||||  |:|| |:||||||||||
Db        198 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 139

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:|||||||  :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        138 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 84

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||  :|||||||  ||||||||||| ||||||||
Db         83 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 25

Qy       4344 AAATGCTTATTAAACTTCCCTTAG 4367
             ||||||||||  ||      | |||
Db          24 AAATGCTTATAGAAGGATGCCTAG 1


RESULT 208
US-10-301-480-998283/c
; Sequence 998283, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 998283
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-998283

  Query Match             6.5%;  Score 488.8;  DB 1;  Length 618;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  535;  Conservative   38;  Mismatches   43;  Indels    8;  Gaps    4;
```

```
Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             |||||||||  :|||| ||||||||||| ||||||||||||:| ||||||||| |||||||
Db        617 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 558

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| ||||||||||||| ||| ||||
Db        557 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 498

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||
Db        497 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 439

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||||| |||||:|||||||| ||||||||||||||||||||||||:
Db        438 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 379

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:|||||||||||| :|||
Db        378 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 319

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             ||  || ||||||||||||||||||: |||||||:|||||||||||||| ||||:|:||||
Db        318 YTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 259

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||: ||||||| | ||||||||||::|||||||:|:|||||||||||||||||||||||:
Db        258 CTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 199

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:||||| |:|| |:||||||||||
Db        198 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 139

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:||||||| :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        138 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 84

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||| :|||||||  |||||||||||| ||||||||
Db         83 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 25

Qy       4344 AAATGCTTATTAAACTTCCCTTAG 4367
             |||||||||  ||     | |||
Db         24 AAATGCTTATAGAAGGATGCCTAG 1
```

RESULT 209
US-09-925-065A-309262/c
; Sequence 309262, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 309262
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-925-065A-309262

Query Match 6.5%; Score 488.4; DB 1; Length 614;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 531; Conservative 38; Mismatches 37; Indels 8; Gaps 4;

```
Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             |||||||||  :|||| ||||||||||| ||||||||||||:| ||||||||| |||||||
Db        613 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 554

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| ||||||||||||| ||| ||||
Db        553 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 494

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||
Db        493 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 435

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||||| |||||:|||||||| ||||||||||||||||||||||||:
Db        434 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 375

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||| ||||:||||:|||||||||||| :|||
Db        374 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 315

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             ||  || ||||||||||||||||||: |||||||:|||||||||||||| ||||:|:||||
Db        314 YTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 255

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||: ||||||| | ||||||||||::|||||||:|:|||||||||||||||||||||||:
Db        254 CTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 195

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:||||| |:|| |:||||||||||
Db        194 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 135

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:||||||| :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        134 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 80

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :|||||||||| :|||||||  |||||||||||| ||||||||
Db         79 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 21

Qy       4344 AAATGCTTATTAAA 4357
             |||||||||  ||
Db         20 AAATGCTTATAGAA 7
```

RESULT 210
US-09-925-065A-309262/c
; Sequence 309262, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 309262
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-309262

  Query Match             6.5%;  Score 488.4;  DB 1;  Length 614;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  531;  Conservative   38;  Mismatches   37;  Indels    8;  Gaps    4;

Qy       3745 CCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAAT 3804
             ||||||||| :|||| |||||||||| ||||||||||:| |||||||| ||||||||
Db        613 CCAGTGCTGTGACTGCGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAAT 554

Qy       3805 GAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGG 3864
             ||||||||||||:||:|||||||||||||:||||| |||||||||||||| ||| ||||
Db        553 GAAGAAAAGATAGAACATAACTGTCAACAAATAATTGCTCAAACCTATGCCGCTCAAGGG 494

Qy       3865 GACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTC 3924
             ||||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||
Db        493 GACCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTC 435

Qy       3925 CTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT-GCAGTGGTCAGTGATAATGGAATAY 3983
             |||||||||||||||||| |||||:|||||| |||||||||||||||||||||||||:
Db        434 CTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGGCAGTGGTCAGTGATAATGGAATAC 375

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
             |||||||||||||||||||||||||||||||| ||||:||||:||||||||||||| :|||
Db        374 TTGAAAGTAATCCCCTCACTCCAGGAACTACTGCTCAGCTGGCAGAACTAATAGCTCTCA 315

Qy       4044 YTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATG 4103
             ||  || ||||||||||||||||: |||||||:||||||||||||| ||||:|:||||
Db        314 YTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGCCTCTAAGTATG 255

Qy       4104 CTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYG 4163
             ||: ||||||| | |||||||||::||||||||:|:|||||||||||||||||||||:
Db        254 CTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCA 195

Qy       4164 AGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAAC 4223
             ||:|||||||  |||:|||||||||||| |||||||:||||| |:|| |:|||||||||
Db        194 AGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCCTGTACAGAAAC 135

Qy       4224 CTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATAS 4283
             |||:|||||||| :|||||||||||||: |||||||||     |||||:|||||||| ||:
Db        134 CTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGAAAGGGAAGTAG 80

Qy       4284 AAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAG 4343
             || :|||:|||||||  :||||||||| :|||||||  |||||||||||| ||||||||
Db          79 AAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGACCTTCCATTAG 21

Qy       4344 AAATGCTTATTAAA 4357
             ||||||||||  ||
Db          20 AAATGCTTATAGAA 7


RESULT 211
US-10-301-480-228755/c
; Sequence 228755, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 228755
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-228755

  Query Match             6.5%;  Score 487;  DB 1;  Length 566;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  516;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             | |||||||||||||||||||| ||||| ||||||||||||||||:|||:||||||||
Db        558 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 499

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
             ||||||| ||||||| |||| |||||||||||||||||||||||| | |:|||||
Db        498 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGAGTTTATCTAAGG 439

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
             |||||||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||
Db        438 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 379

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        378 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 319

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
             ||||||| |||  |||||||||||| |||||||: || | ||   ||||||||||
Db        318 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCAYGCATATTG--TTGTCTGATGGGG 261

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
             |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||| |||:
Db        260 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 201

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
             |||||||| |||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        200 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 141

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
             ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        140 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 81

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
             |||| |||||||||||||||||||||||||||||||||| |||||||||| ||||||:|||| |
Db         80 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 21

Qy       6173 TTTCAATCCCTGTACCTGAA 6192
             |||||||||||||| | |||
Db         20 TTTCAATCCCTGTAGCCGAA 1


RESULT 212
US-10-301-480-228756/c
; Sequence 228756, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 228756
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-228756
```

```
  Query Match             6.5%;  Score 487;  DB 1;  Length 566;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  516;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             | |||||||||||||||||||||| ||||| |||||||||||||||:|||:||||||||
Db        558 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 499

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
             ||||||| ||||||| |||| |||||||||||||||||||:||||||||| | |:|||||
Db        498 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCRTATAGGAGTTTATCTAAGG 439

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
             |||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db        438 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 379

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        378 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 319

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
             ||||||| |||  |||||||||||||| |||||||||| || | ||   ||||||||||||
Db        318 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCACGCATATTG--TTGTCTGATGGGG 261

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
             |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||| |||:
Db        260 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 201

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
             |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 141

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
             ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 81

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
             |||| |||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db         80 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 21

Qy       6173 TTTCAATCCCTGTACCTGAA 6192
             |||||||||||||| | |||
Db         20 TTTCAATCCCTGTAGCCGAA 1


RESULT 213
US-10-301-480-842164/c
; Sequence 842164, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 842164
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-842164

  Query Match             6.5%;  Score 487;  DB 1;  Length 566;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  516;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             | |||||||||||||||||||||| ||||| |||||||||||||||:|||:||||||||
Db        558 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 499

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
             ||||||| ||||||| |||| |||||||||||||||||||:||||||||| | |:|||||
Db        498 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCRTATAGGAGTTTATCTAAGG 439

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
             |||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db        438 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 379

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        378 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 319

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
             ||||||| |||  |||||||||||||| |||||||||: || | ||   ||||||||||||
Db        318 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCAYGCATATTG--TTGTCTGATGGGG 261

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
             |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||| |||:
Db        260 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 201

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
             |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 141

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
             ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 81

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
             |||| |||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db         80 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 21

Qy       6173 TTTCAATCCCTGTACCTGAA 6192
             |||||||||||||| | |||
Db         20 TTTCAATCCCTGTAGCCGAA 1


RESULT 214
US-10-301-480-842165/c
; Sequence 842165, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 842165
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-842165

  Query Match             6.5%;  Score 487;  DB 1;  Length 566;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  516;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             | |||||||||||||||||||||| ||||| |||||||||||||||:|||:||||||||
Db        558 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 499

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
```

```
                |||||||  |||||||  ||||  ||||||||||||||||||||:|||||||||  | |:|||||
Db        498 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCRTATAGGAGTTTATCTAAGG 439

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
                |||||||||||||||||||||||||||||||||  |||||||  ||||||  ||||||||||||
Db        438 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 379

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
                |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||  |||
Db        378 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 319

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
                |||||||  |||   ||||||||||||  ||||||||||  ||  |  ||   ||||||||||||
Db        318 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCACGCATATTG--TTGTCTGATGGGG 261

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
                |||||||||||||||||||||||||||||||||  |||||  ||||||||||||||||  |||:
Db        260 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 201

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
                ||||||||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 141

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
                |||  ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 81

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
                ||||  ||||||||||||||||||||||||||||||  ||||||||||||  ||||||:||||  |
Db          80 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 21

Qy       6173 TTTCAATCCCTGTACCTGAA 6192
                |||||||||||||||  | |||
Db          20 TTTCAATCCCTGTAGCCGAA 1


RESULT 215
US-09-925-065A-195740/c
; Sequence 195740, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195740
;   LENGTH: 595
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-195740

  Query Match             6.5%;  Score 486.2;  DB 1;  Length 595;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  518;  Conservative   37;  Mismatches   32;  Indels   12;  Gaps    3;

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
                |||||||||||||  |||||||:||  ||::|||||||||||||||||||||||||||||||
Db        595 TTACTGTGTGGACCTCTCATGATGTAAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 536

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
                ||||||||||  ||||  |||||||:|||||||||||||||||||:||||||||||||:||
Db        535 TGTCAGACAACAGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAAC 476

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
                ||  |||||||||||||||||||||||:|  |||||||||||||||||||||||||||||:
Db        475 TGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAG 416

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
                ||:|||||||||||||  :| |||  ||  ||||  ||  |||  ||  ||||||||||||:||||:|
Db        415 AACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGACCTTCTAGAGG 356

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA--------GTTCCTTTG 3929
                ||||:||||||||||:||||  ||||||||||  ||||||||||        |||||||  |
Db        355 TTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTTTGGTTCCTTGG 297

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
                ||||||||||||||||  |||||:|||||  ||||||||||||||||||||||||:||||||
Db        296 TAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGGAATACTTGAAA 237

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
                |||||||||  |||||||  ||||||  |||:||||:||||||  |||||:  |:|||:|  |||
Db        236 GTAATCCCCTTACTCCAGAAACTAGCGCTCAGCTGGCAGAATTAATASGCCTCACTCGGG 177

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
                |||||  ||||   |||||:|||||||:||||  ||||||||||||||:|:||||||:|  |
Db        176 CACTAAAATT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAAGTATGCTTATC 120

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
                |||||  ||||||||  |  ::|||||||||:|:||||||||||||||||||||||:|||:|||
Db        119 TAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAA 60

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
                |||||||||:|||||||||||||||||||:|||||||:||||:|||||||||||||:|
Db         59 CACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAA 1


RESULT 216
US-09-925-065A-195740/c
; Sequence 195740, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195740
;   LENGTH: 595
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-195740

  Query Match             6.5%;  Score 486.2;  DB 1;  Length 595;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  518;  Conservative   37;  Mismatches   32;  Indels   12;  Gaps    3;

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
                |||||||||||||  |||||||:||  ||::|||||||||||||||||||||||||||||||
Db        595 TTACTGTGTGGACCTCTCATGATGTAAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 536
```

```
Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
             ||||||||||| |||| |||||||:|||||||||||||||||||:||||||||||||:||
Db        535 TGTCAGACAACAGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAAC 476

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
             ||  ||||||||||||||||||||:| |||||||||||||||||||||||||||||||:
Db        475 TGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAG 416

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
             ||:|||||||||||||| :| ||| || |||| || ||| || ||||||||||:||||:|
Db        415 AACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGACCTTCTAGAGG 356

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA--------GTTCCTTTG 3929
             ||||:|||||||||||:|||| |||||||||| ||||||||||        |||||||.|
Db        355 TTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTTTGGTTCCTTGG 297

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
             |||||||||||||||| |||||:||||| ||||||||||||||||||||||||:||||||
Db        296 TAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGGAATACTTGAAA 237

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
             |||||||||| ||||||| |||||| |||:||||:|||||| |||||: |:|||:| |||
Db        236 GTAATCCCCTTACTCCAGAAACTAGCGCTCAGCTGGCAGAATTAATASGCCTCACTCGGG 177

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
             ||||| ||||    |||||:||||||||:|||| |||||||||||||:|:|||||||:| |
Db        176 CACTAAAATT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAAGTATGCTTATC 120

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
             ||||| |||||||| |  ::|||||||:|:||||||||||||||||||||:|||:|||
Db        119 TAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAA 60

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
             |||||||||:|||||||||||||||||||:||||||:||||:|||||||||||||:|
Db          59 CACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAA 1


RESULT 217
US-10-301-480-284876/c
; Sequence 284876, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 284876
;   LENGTH: 595
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-284876

  Query Match             6.5%;  Score 486.2;  DB 1;  Length 595;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  518;  Conservative   37;  Mismatches   32;  Indels   12;  Gaps    3;

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
             |||||||||||||| ||||||||:|| ||::|||||||||||||||||||||||||||||
Db        595 TTACTGTGTGGACCTCTCATGATGTAAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 536

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
             |||||||||||| |||||||:|||||||||||||||||||:||||||||||||:||
Db        535 TGTCAGACAACAGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAAC 476

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
             ||  ||||||||||||||||||||:| |||||||||||||||||||||||||||||||:
Db        475 TGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAG 416

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
             ||:|||||||||||||| :| ||| || |||| || ||| || ||||||||||:||||:|
Db        415 AACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGACCTTCTAGAGG 356

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA--------GTTCCTTTG 3929
             ||||:|||||||||||:|||| |||||||||| ||||||||||        |||||||.|
Db        355 TTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTTTGGTTCCTTGG 297

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
             |||||||||||||||| |||||:||||| ||||||||||||||||||||||||:||||||
Db        296 TAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGGAATACTTGAAA 237

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
             |||||||||| ||||||| |||||| |||:||||:|||||| |||||: |:|||:| |||
Db        236 GTAATCCCCTTACTCCAGAAACTAGCGCTCAGCTGGCAGAATTAATASGCCTCACTCGGG 177

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
             ||||| ||||    |||||:||||||||:|||| |||||||||||||:|:|||||||:| |
Db        176 CACTAAAATT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAAGTATGCTTATC 120

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
             ||||| |||||||| |  ::|||||||:|:||||||||||||||||||||:|||:|||
Db        119 TAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAA 60

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
             |||||||||:|||||||||||||||||||:||||||:||||:|||||||||||||:|
Db          59 CACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAA 1


RESULT 218
US-10-301-480-898285/c
; Sequence 898285, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 898285
;   LENGTH: 595
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-898285

  Query Match             6.5%;  Score 486.2;  DB 1;  Length 595;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  518;  Conservative   37;  Mismatches   32;  Indels   12;  Gaps    3;

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
             |||||||||||||| ||||||||:|| ||::|||||||||||||||||||||||||||||
Db        595 TTACTGTGTGGACCTCTCATGATGTAAACAGCATACTCACTGCTAAAGGAGACTTGTGGC 536

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
             |||||||||||| |||||||:|||||||||||||||||||:||||||||||||:||
Db        535 TGTCAGACAACAGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCAAC 476

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
             ||  ||||||||||||||||||||:| |||||||||||||||||||||||||||||||:
Db        475 TGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAG 416

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
```

```
                ||:|||||||||||||| :| ||| || |||| || ||| || ||||||||||:||||:|
Db        415 AACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGACCTTCTAGAGG 356

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA--------GTTCCTTTG 3929
                ||||:|||||||||||:|||| |||||||||| ||||||||||        |||||| |
Db        355 TTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTTTGGTTCCTTGG 297

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
                |||||||||||||| |||||:||||| ||||||||||||||||||||||||:||||||
Db        296 TAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGGAATACTTGAAA 237

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
                |||||||||| ||||||| |||||| |||:||||:|||||| |||||: |:|||:| |||
Db        236 GTAATCCCCTTACTCCAGAAACTAGCGCTCAGCTGGCAGAATTAATASGCCTCACTCGGG 177

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
                ||||| ||||   |||||:|||||||:|||| ||||||||||||||:|:|||||||:| |
Db        176 CACTAAAATT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAAGTATGCTTATC 120

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
                ||||| |||||||| |  ::|||||||:|:||||||||||||||||||:|||:|||
Db        119 TAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGGAA 60

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
                |||||||||:|||||||||||||||||||:|||||||:||||:|||||||||||||:|
Db          59 CACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAA 1
```

RESULT 219
US-09-925-065A-928578
; Sequence 928578, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928578
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-928578

Query Match 6.5%; Score 486.2; DB 1; Length 681;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 590; Conservative 2; Mismatches 81; Indels 14; Gaps 5;

```
Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
                |||||||||||||||||| ||||||| |||||||||||||||||| ||||||| |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60

Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
                |||| || |||||||||| |||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                     |||||||||||||| |||||||||| |||||||||  || ||||||||||
Db        121 TAAAAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
                |||||||||||||  |||||||||  ||| |||: |||||||||||||||||||| ||||
Db        181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCYGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
                |||||||||||||||||||||||||||| | ||||| ||||| ||||| ||||||||| |
Db        241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
                ||| ||  ||| |||||||||||||||||||||||||||| |||||| |||| | |||||
Db        301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
                |||||||||||||||||||||||||||||||||||||||||||||  |||||||| |||
Db        361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
                |||||||||||| |||||||||| | || |||||||||||||||||| |||||| | |
Db        421 CCCATGCCATCACTCTCACAGAGCGCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                  |||||||||||||||||  | | ||||| |||||||||| ||||||||||||||||
Db        481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
                ||||||| ||||||||   | | | ||| | || ||| ||||||||||||||||| ||:
Db        541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
                |||||||||||||||| |||   ||| |  ||| |   |||||||||| |||| |||| |
Db        597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGG 2327
                |  |||||||||  || | ||||||||
Db        655 GTCTGAAAGCCCTGCTCCTTTGTTAGG 681
```

RESULT 220
US-09-925-065A-928578
; Sequence 928578, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928578
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-928578

Query Match 6.5%; Score 486.2; DB 1; Length 681;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 590; Conservative 2; Mismatches 81; Indels 14; Gaps 5;

```
Qy       1649 AAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGG 1708
                |||||||||||||||||| ||||||| |||||||||||||||||| ||||||| |
Db          1 AAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTAG 60
```

```
Qy       1709 CAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATT 1768
              |||| || |||||||||||| |||||||||||||||||||| |||||||||||||| |||||
Db         61 CAACATCCGTTTTTTATAAGAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACAGGATT 120

Qy       1769 -----AAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
                   |||||||||||||| |||||||||| |||||||||||  || ||||||||||
Db        121 TAAAAAAAAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGATTTTGGAGGCT 180

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
              |||||||||||||  |||||||||  ||| |||: |||||||||||||||||||| ||||
Db        181 TTGGAAAAGGGAAAGCCTGGGCAAACAGAACGCCYGATAGGGCTTGCTTCCAGTGTGGTC 240

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
              |||||||||||||||||||||||||||| | ||||| ||||| ||||| ||||||||| |
Db        241 TACAAGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCTGCCCCCTCGTCCATGTC 300

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
              ||| ||  ||| |||||||||||||||||||||||||||| |||||| |||| | |||||
Db        301 CCTCATGACAATGGAATCACTGGAAGGCCCACTGCCCCAGAGGACAATGGTCCTCTGAGT 360

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||| |||
Db        361 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCACCAG 420

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
              |||||||||||||| |||||||||| | || |||||||||||||||||| |||||| | |
Db        421 CCCATGCCATCACTCTCACAGAGCGCCGGCTATGCTTGACCATTGAGGACCAGGAGGTTA 480

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                |||||||||||||||||  | | ||||| |||||||||||| ||||||||||||||||||
Db        481 ACTGTCTCCTGGACACTGGCACAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGT 540

Qy       2181 CCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTY 2240
              ||||||| ||||||||   | |  | |||  | || ||| |||||||||||||||| ||:
Db        541 CCTCCAGGTCTGTCAC---TGTCTGAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTC 596

Qy       2241 TCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTAT 2300
              |||||||||||||||| |||   ||| |  ||| |   |||||||||| |||| |||| |
Db        597 TCCCAGCCACTAAGTTGTGACTGGGGAACTTTTCT--CTTTCACATGCCTTTCAAATTTT 654

Qy       2301 GCTTGAAAGCCCCACTACCTTGTTAGG 2327
              |  |||||||||  || | ||||||||
Db        655 GTCTGAAAGCCCTGCTCCTTTGTTAGG 681
```

```
RESULT 221
US-09-925-065A-132544/c
; Sequence 132544, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132544
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-09-925-065A-132544

  Query Match             6.5%;  Score 486;  DB 1;  Length 570;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  515;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
              | |||||||||||||||||||||||| ||||| ||||||||||||||||:|||:||||||||
Db         557 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 498

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
              ||||||| ||||||| |||| |||||||||||||||||||||||||||| | |:|||||
Db         497 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGAGTTTATCTAAGG 438

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
              |||||||||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db         437 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 378

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         377 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 318

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
              ||||||| |||  |||||||||||| |||||||||: || | ||   ||||||||||||
Db         317 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCAYGCATATTG--TTGTCTGATGGGG 260

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
              |||||||||||||||||||||||||||||||| ||||| |||||||||||||||| |||:
Db         259 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 200

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
              |||||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db         199 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 140

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
              ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         139 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 80

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
              |||| |||||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db          79 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 20

Qy       6173 TTTCAATCCCTGTACCTGA 6191
              ||||||||||||||| | ||
Db          19 TTTCAATCCCTGTAGCCGA 1
```

```
RESULT 222
US-09-925-065A-132544/c
; Sequence 132544, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132544
;   LENGTH: 570
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-925-065A-132544

  Query Match             6.5%;  Score 486;  DB 1;  Length 570;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  515;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
              | |||||||||||||||||||||| ||||| ||||||||||||||||:|||:||||||||
Db        557 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 498

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
              ||||||| ||||||| |||| |||||||||||||||||||||||| | |:|||||
Db        497 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGAGTTTATCTAAGG 438

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
              ||||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db        437 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 378

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
              ||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        377 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 318

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
              ||||||| |||  |||||||||||| |||||||||: || | ||   ||||||||||||
Db        317 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCAYGCATATTG--TTGTCTGATGGGG 260

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
              |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||| |||:
Db        259 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 200

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
              |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 140

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
              ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 80

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
              |||| |||||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db          79 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 20

Qy       6173 TTTCAATCCCTGTACCTGA 6191
              ||||||||||||||| | ||
Db          19 TTTCAATCCCTGTAGCCGA 1


RESULT 223
US-09-925-065A-132545/c
; Sequence 132545, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132545
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-132545

  Query Match             6.5%;  Score 486;  DB 1;  Length 570;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  515;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
              | |||||||||||||||||||||| ||||| ||||||||||||||||:|||:||||||||
Db        557 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 498

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
              ||||||| ||||||| |||| |||||||||||||||||||||||| | |:|||||
Db        497 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCRTATAGGAGTTTATCTAAGG 438

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
              ||||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db        437 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 378

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
              ||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        377 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 318

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
              ||||||| |||  |||||||||||| |||||||||: || | ||   ||||||||||||
Db        317 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCACGCATATTG--TTGTCTGATGGGG 260

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
              |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||| |||:
Db        259 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 200

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
              |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 140

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
              ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 80

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
              |||| |||||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db          79 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 20

Qy       6173 TTTCAATCCCTGTACCTGA 6191
              ||||||||||||||| | ||
Db          19 TTTCAATCCCTGTAGCCGA 1


RESULT 224
US-09-925-065A-132545/c
; Sequence 132545, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132545
```

```
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-132545

  Query Match             6.5%;  Score 486;  DB 1;  Length 570;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  515;  Conservative    6;  Mismatches   36;  Indels    2;  Gaps    1;

Qy       5633 CTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTC 5692
             | |||||||||||||||||||||| ||||| ||||||||||||||||:|||:||||||||
Db        557 CACTCACTGCACCCCCTCCATGCCACTGTACGACCAGTAGCTCCCCTTACCAAGAGTTTC 498

Qy       5693 TATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGG 5752
             ||||||| ||||||| |||| ||||||||||||||||||:||||||||| | |:|||||
Db        497 TATGGAGGATGCAGCTTCCCAGAAATATTGATGCCCCATCRTATAGGAGTTTATCTAAGG 438

Qy       5753 GAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTC 5812
             |||||||||||||||||||||||||||||||||| ||||||| |||||| ||||||||||||
Db        437 GAACCCCCACCTTCACTGCCCACACCCATATGTCCCGCAATTGCTATAACTCTGCCACTC 378

Qy       5813 TTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTG 5872
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db        377 TTTGCATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTTCTG 318

Qy       5873 GAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGG 5932
             ||||||| |||  |||||||||||||| |||||||||| || | ||   ||||||||||||
Db        317 GAGGACTGGGAACCACTGTCTGTTGGGCTTACTTCACGCATATTG--TTGTCTGATGGGG 260

Qy       5933 GTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCS 5992
             |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||| |||:
Db        259 GTGGAGTTCAAGATCAGGCAAGAGAAAAACACGTAAAGGAAGTAATCTCCCAACTGACCC 200

Qy       5993 GGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 6052
             |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GGGTACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCC 140

Qy       6053 TCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 6112
             ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TCCATACCCATACTTGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGG 80

Qy       6113 TCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATG 6172
             |||| |||||||||||||||||||||||||| |||||||||||| ||||||:|||| |
Db          79 TCTCAGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCACACA 20

Qy       6173 TTTCAATCCCTGTACCTGA 6191
             ||||||||||||| | ||
Db          19 TTTCAATCCCTGTAGCCGA 1


RESULT 225
US-09-925-065A-955763/c
; Sequence 955763, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 955763
;   LENGTH: 615
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-955763

  Query Match             6.5%;  Score 484.8;  DB 1;  Length 615;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  523;  Conservative   29;  Mismatches   53;  Indels   10;  Gaps    2;

Qy       4027 AGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATA 4086
             |||||||||||||:|||:| |||||||||||||||||||:||||||||:|||||||||
Db        615 AGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATA 556

Qy       4087 TACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAG 4146
             ||||| ||||:|:||||||:|||||||  ||||||||||:|::|||||||||:|:||||||
Db        555 TACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAAAG 496

Qy       4147 GGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATT 4206
             |||||||||||||||||: ||:|||||||||||||:|||||| |||||||||||||| ||||||
Db        495 GGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTATT 436

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT-------C 4259
             |:||||:|||||||||||||:||||||||:|||||||||||||:||||||||
Db        435 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAAAAA 376

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
             | |||  ||||:|||||||| ||:|||:|||:||||||||:|| |||||||:|||||||
Db        375 AAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAGAG 316

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
             |  |||| |||||||||||||||||||||||||||||||  ||   ||| |||| | |||||| |
Db        315 CCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAACC 256

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
             || ||| |||||||||||||||||||||||||||| ||||| ||||||| |||||||| |||
Db        255 CCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAGAG 196

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             |    | ||| |||||||| ||| || |||||||  ||||||  ||||||||||||||||||
Db        195 GACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCCTG 136

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             || ||| ||||| |||||||||||||||||||| | |||||||||||||||||||| ||||
Db        135 CAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGATA 76

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||| | ||||:||||| |||| ||||||||||||||||||||||||||||||||||||||
Db         75 GCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAAT 16

Qy       4617 ARATAKTCAGGGCCT 4631
             |:|||:|||||||||
Db         15 AGATAGTCAGGGCCT 1


RESULT 226
US-09-925-065A-955763/c
; Sequence 955763, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 955763
;   LENGTH: 615
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-955763

  Query Match             6.5%;  Score 484.8;  DB 1;  Length 615;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  523;  Conservative   29;  Mismatches   53;  Indels   10;  Gaps    2;

Qy       4027 AGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATA 4086
             ||||||||||||||:|||:| ||||||||||||||||||||:|||||||:|||||||||
Db        615 AGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATA 556

Qy       4087 TACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAG 4146
             ||||| |||||:|:||||||:|||||||  ||||||||||:|::|||||||:|:||||||
Db        555 TACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAAAG 496

Qy       4147 GGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATT 4206
             |||||||||||||||: ||:||||||||||||:|||||| |||||||||||| ||||||
Db        495 GGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTATT 436

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT-------C 4259
             |:||||:|||||||||||||:|||||||:|||||||||||||:|||||||
Db        435 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAAAA 376

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
             | |||  ||||:|||||||| ||:|||:|||:||||||||:|| |||||||:|||||||
Db        375 AAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAGAG 316

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
             |  |||| ||||||||||||||||||||||||||  ||   ||| |||| | |||||| |
Db        315 CCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAACC 256

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
             || ||| |||||||||||||||||||||||||| ||||| ||||||| |||||||| |||
Db        255 CCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAGAG 196

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             |    | ||| |||||||| ||| || |||||||  ||||||  ||||||||||||||||
Db        195 GACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCCTG 136

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             || ||| ||||| |||||||||||||||||| | ||||||||||||||||||||| ||||
Db        135 CAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGATA 76

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||| | ||||:||||| |||| ||||||||||||||||||||||||||||||||||||
Db          75 GCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAAT 16

Qy       4617 ARATAKTCAGGGCCT 4631
             |:|||:|||||||||
Db          15 AGATAGTCAGGGCCT 1


RESULT 227
US-09-925-065A-955838/c
; Sequence 955838, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 955838
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-955838

  Query Match             6.5%;  Score 483.8;  DB 1;  Length 614;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  522;  Conservative   29;  Mismatches   53;  Indels   10;  Gaps    2;

Qy       4027 AGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATA 4086
             ||||||||||||||:|||:| ||||||||||||||||||||:|||||||:|||||||||
Db        614 AGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATA 555

Qy       4087 TACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAG 4146
             ||||| ||||:|:|||||||:|||||||  ||||||||||:|::|||||||:|:||||||
Db        554 TACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAAAG 495

Qy       4147 GGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATT 4206
             |||||||||||||||: ||:||||||||||||:|||||| |||||||||||| ||||||
Db        494 GGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTATT 435

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT-------C 4259
             |:||||:|||||||||||||:|||||||:|||||||||||||:|||||||
Db        434 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAAAA 375

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
             | |||  ||||:|||||||| ||:|||:|||:||||||||:|| |||||||:|||||||
Db        374 AAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAGAG 315

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
             |  |||| ||||||||||||||||||||||||||  ||   ||| |||| | |||||| |
Db        314 CCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAACC 255

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
             || ||| |||||||||||||||||||||||||| ||||| ||||||| |||||||| |||
Db        254 CCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAGAG 195

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             |    | ||| |||||||| ||| || |||||||  ||||||  ||||||||||||||||
Db        194 GACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCCTG 135

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             || ||| ||||| |||||||||||||||||| | ||||||||||||||||||||| ||||
Db        134 CAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGATA 75

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||| | ||||:||||| |||| ||||||||||||||||||||||||||||||||||||
Db          74 GCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAAT 15

Qy       4617 ARATAKTCAGGGCC 4630
             |:|||:|||||||||
Db          14 AGATAGTCAGGGCC 1


RESULT 228
US-09-925-065A-955838/c
; Sequence 955838, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
```

```
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 955838
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-955838

  Query Match             6.5%;  Score 483.8;  DB 1;  Length 614;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  522;  Conservative   29;  Mismatches   53;  Indels   10;  Gaps    2;

Qy       4027 AGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATA 4086
             |||||||||||||:|||:| ||||||||||||||||||||:||||||||:||||||||
Db        614 AGAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATA 555

Qy       4087 TACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAG 4146
             ||||| ||||:|:||||||:|||||||  ||||||||||:|::||||||||:|:||||||
Db        554 TACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCAYGCAGCAATATGGAGAGAAAG 495

Qy       4147 GGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATT 4206
             |||||||||||||||: ||:||||||||||||:|||||| |||||||||||||| ||||||
Db        494 GGAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTATT 435

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT-------C 4259
             |:||||:|||||||||||||:||||||||:|||||||||||||||:||||||||
Db        434 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAAAAA 375

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
             | |||  ||||:|||||||| ||:|||:|||:|||||||||:|| |||||||:|||||||
Db        374 AAAAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAGAG 315

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
             |  |||| |||||||||||||||||||||||||||  ||   ||| |||| | |||||| |
Db        314 CCACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAACC 255

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
             || ||| |||||||||||||||||||||||||| ||||| ||||||| |||||||| |||
Db        254 CCCTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAGAG 195

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             |   | ||| |||||||| ||| || ||||||| ||||||  ||||||||||||||||
Db        194 GACATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCCTG 135

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             || ||| ||||| |||||||||||||||||||| | |||||||||||||||||||| ||||
Db        134 CAGCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGATA 75

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||| | ||||:||||| |||| ||||||||||||||||||||||||||||||||||||
Db          74 GCAACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAAT 15

Qy       4617 ARATAKTCAGGGCC 4630
             |:|||:||||||||
Db          14 AGATAGTCAGGGCC 1


RESULT 229
US-09-925-065A-409227
; Sequence 409227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 409227
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-409227

  Query Match             6.5%;  Score 483.4;  DB 1;  Length 593;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  514;  Conservative   29;  Mismatches   42;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             ||||||||||||||||||||||||||||||||||||||||| |||||||:|||||||||||
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||||||||||||||| ||||||||||||:||||:||||||||
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| ||||||||||||||||||||||:|||||||||:|||||||||||||||
Db        121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |       |||: ||||||| |||||||||||||::|||||||||:|:|||||||
Db        181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGASAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             |||||||||||||:|||: ||||||||||:|||||||||||||||||| ||:||||| |
Db        241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCA 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:|||||||||||||:||||||| :|||||||||||||: |||||||| |||| |
Db        301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||||| |:|| :|||:  |||||| :||||||||||:|||||||||||||||
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||| |||||||||  ||   ||| |||||| ||||||| || ||| |
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             |||||||||||||||||||||||||||||||||| |||||||||||||||| |||     |||
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGC 4497
             || |||||||| ||| || ||||||| |||||||||||||||||||| ||||||
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGC 593


RESULT 230
US-09-925-065A-409227
```

```
; Sequence 409227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 409227
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-409227

  Query Match             6.5%;  Score 483.4;  DB 1;  Length 593;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  514;  Conservative   29;  Mismatches   42;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
Db        181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGASAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
Db        241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCA 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
Db        301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
Db        361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
Db        421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
Db        481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGC 4497
Db        541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGC 593

RESULT 231
US-10-301-480-373537/c
; Sequence 373537, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 373537
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-373537

  Query Match             6.5%;  Score 483.3;  DB 1;  Length 583;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  499;  Conservative   56;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
Db        583 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 524

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
Db        523 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 464

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
Db        463 AAACCTATCTAGCCTATACTGGCTTATCCTCATCCCAAAACACTAAAGCAACTAAGAGGG 404

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
Db        403 TTCCTTGGCATAAAAGGCTTCTGCCGAATATGGATT-CCCAGGTACAGCGAAATAGCCAG 345

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
Db        344 GCCATTATATACACTAATTAAGGAAACTCASAAAGCCAATACCCATTTAGTAAGATGGAT 285

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
Db        284 ACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTG 227

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
Db        226 TTAAGCTTGCCAACAGTGCTAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGC 169

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
Db        168 TCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAA 109

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
Db        108 GGAAACTGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGTGGCAGCAGTAGC 49

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
Db         48 AGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGATCTTAC 1

RESULT 232
US-10-301-480-986946/c
; Sequence 986946, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 986946
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-986946

  Query Match             6.5%;  Score 483.3;  DB 1;  Length 583;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  499;  Conservative   56;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:|||||||||| |:|||||||||||||:|||||||| ||:|:|:|||||:|||||| 
Db        583 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 524

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||||:||||:||||||||||||||||:|||| |||||  || 
Db        523 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 464

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
               ||: :||| |||||||||||||||||||||||:|||||| |||||||||||||||:|:: 
Db        463 AAACCTATCTAGCCTATACTGGCTTATCCTCATCCCAAAACACTAAAGCAACTAAGAGGG 404

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
             ||||||||||||| ||||:|||||||||||:||||||| |||||||:  ||:|||||||||| 
Db        403 TTCCTTGGCATAAAAGGCTTCTGCCGAATATGGATT-CCCAGGTACAGCGAAATAGCCAG 345

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
             |:|||||:|||||:|||||||||||||||||:|||||||||||||||||||:||||||||: 
Db        344 GCCATTATATACACTAATTAAGGAAACTCASAAAGCCAATACCCATTTAGTAAGATGGAT 285

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
             |:||||||:::|||||||||||||||||||  ||  | | |||  ||||||| :||||||| 
Db        284 ACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTG 227

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
             |||||:||||||||:| || |||||||||:||:|||::|||||| |||||||||:||:||| 
Db        226 TTAAGCTTGCCAACAGTGCTAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGC 169

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
             |||:|||||||||||||:|| :|||||: |||||||||||::||||::|| |||:||| 
Db        168 TCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAA 109

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
             |||||:|||||||||| ||||||||||:|:||||||||||: |||||||  |||||||| 
Db        108 GGAAACTGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGTGGCAGCAGTAGC 49

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
             |||::|||||||||||||||||||||||| ||||||||:|||||||||| 
Db         48 AGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGATCTTAC 1


RESULT 233
US-09-925-065A-409228
; Sequence 409228, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 409228
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-409228

  Query Match             6.4%;  Score 483;  DB 1;  Length 593;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  513;  Conservative   30;  Mismatches   42;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             |||||||||||||||||||||||||||||||||||||| ||||||:|||||||||||||||| 
Db          1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:|||||||||||||||||||||||| ||||||||||||:||||:|||||||| 
Db         61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| |||||||||||||||||||||:||||||||:|||||||||||||||| 
Db         121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |       |||: ||||||| |||||||||||::|||||||||:|:||||||| 
Db         181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             ||||||||||||||:|||: |||||||||||:|||||||||||||||||| ||:||||| : 
Db         241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCR 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:|||||||||||:|||||||| :|||||||||||||: ||||||||| |||| | 
Db         301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||||| |:|| :|||:  |||||| :|||||||||:|||||||||||||||| 
Db         361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||||| ||||||||||  ||  ||| |||||| ||||||| || ||| | 
Db         421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             |||||||||||||||||||||||||||||||| |||||||||||||||| |||     ||| 
Db         481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGC 4497
             || |||||||| ||| || ||||||| |||||||||||||||||||||||| |||||| 
Db         541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGC 593


RESULT 234
US-09-925-065A-409228
; Sequence 409228, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 409228
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-409228

  Query Match             6.4%;  Score 483;  DB 1;  Length 593;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  513;  Conservative   30;  Mismatches   42;  Indels    8;  Gaps    2;

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
             ||||||||||||||||||||||||||||||| ||||||:||||||||||||||||||||||
Db             1 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTGA 60

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
             ||||||||||:||||||||||||||||||||| ||||||||||:||||:|||||||||||
Db            61 TAATGGAATACTTGAAAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACT 120

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||||:|||:| ||||||||||||||||||||:||||||||:|||||||||||||||||
Db           121 AATAGCCCTCACTCGGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGA 180

Qy       4093 CTCTRARTA-----TGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
             | || |: |      |||: ||||||| ||||||||||::||||||||:|:|||||||||
Db           181 CACTTAAGACTAGGGGCTTGCCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGG 240

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
             ||||||||||||||:|||: ||||||||||:||||||||||||||||| ||:|||||| :
Db           241 GAATTCCTAACTTCTGAGGCAACACCTATCAAACATCAGGAAGCCATTAAGAGATTATCR 300

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
             :||||:|||||||||||||:|||||||  :|||||||||||||: ||||||||| |||| |
Db           301 TTGGCTGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCAGAAAGAA 360

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
             |||:|||||||||| |:|| :|||:  |||||| :||||||||||:|||||||||||||||
Db           361 AAGGAAAGGGAAACAGAATGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGG 420

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
             |||||||||||||||| ||||||||||   ||    ||| |||||| ||||||| || ||| |
Db           421 CAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATACCCTCCAG 480

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGT 4444
             |||||||||||||||||||||||||||||||| ||||||||||||||||  |||     |||
Db           481 GAAACCAAGCCCCAGTACTCAGCAGGAGAAATAGAATGGGGAACCTCAAGAGCACATAGT 540

Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGC 4497
             || |||||||| ||| || ||||||| |||||||||||||||||| ||||||
Db           541 TTCCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGGAAAATACTTTAGCCTGC 593


RESULT 235
US-09-925-065A-755821/c
; Sequence 755821, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 755821
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-755821

  Query Match             6.4%;  Score 481.9;  DB 1;  Length 631;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  571;  Conservative    8;  Mismatches   50;  Indels    9;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
             |||||||| || ||||||||||||||||||:|||||||||||| ||:||||||||||||
Db           631 AGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG 572

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                 | |       |    |||||||||||| |||||||||| ||||||||||||||||
Db           571 ----ACGCAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACTTGTGTGCTA 516

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
              :||||||||||||||||||:|||  ||: || :|||||| |||||||||||||| |||||
Db           515 GAAGGACTAAGGAAAACTAGGAA-TAAGCCTGTGAATTATTCAATGATGTCCACTATAAC 457

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||| ||||||
Db           456 ACAGGGAAAGGG-AGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAATGAGGA 398

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
              ||| | |||| |||||||||||||||  |||||:|||||||||| ||||||  ||||||||
Db           397 AGCATACCTCCCTGTCACCTGACTCTATTGAAGRCCAACTAATGTTAAAGGATAAGTTTA 338

Qy       1623 TCACTCAGTCAGCTGCAGACATTAGAAAAAA-CTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
              ||||||| |||||||||||||||||||||||| |||||||||||||| ||||||| ||||
Db           337 TCACTCATTCAGCTGCAGACATTAGAAAAAAGCTTCAAAAGTCTGCCTTAGGCCCAGAGC 278

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
              |||||||||||||||||||||||||||||||:|| |||||||||||||||||||| ||
Db           277 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCAGTTTTTTATAATAGAGATCAGGCGG 218

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
              ||||||| |||| |||||||| ||| |||| |||||||||||||||||||||| |||||
Db           217 AGCAGGCAGAACGGGACAAACAGGA-TAAAGAAAAGGCCACCGCTTTAGTCATGGCCCTC 159

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           158 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 99

Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db            98 AGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAAT 39

Qy       1922 AAGCCG-CCCCTTCGTCCATGCCCCTTATTTCAAGGGA 1958
              |||| | |||| || |||||||||||||| ||||||||
Db            38 AAGCTGCCCCCCTCATCCATGCCCCTTATGTCAAGGGA 1

RESULT 236
US-09-925-065A-755821/c
; Sequence 755821, Application US/09925065A
; Publication No. US20050228172A9
```

```
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 755821
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-755821

  Query Match             6.4%;  Score 481.9;  DB 1;  Length 631;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  571;  Conservative    8;  Mismatches   50;  Indels    9;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
             |||||||| || ||||||||||||||||||:||||||||||| ||:|||||||||||||
Db        631 AGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG 572

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                 | |       |    |||||||||||| |||||||||| ||||||||||||||||
Db        571 ----ACGCAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACTTGTGTGCTA 516

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
              :||||||||||||||||||:|||  ||: || :|||||| ||||||||||||||| |||||
Db        515 GAAGGACTAAGGAAAACTAGGAA-TAAGCCTGTGAATTATTCAATGATGTCCACTATAAC 457

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||| ||||||
Db        456 ACAGGGAAAGGG-AGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAATGAGGA 398

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
              ||| | |||| |||||||||||||||| |||||:|||||||| ||||||  ||||||||
Db        397 AGCATACCTCCCTGTCACCTGACTCTATTGAAGRCCAACTAATGTTAAAGGATAAGTTTA 338

Qy       1623 TCACTCAGTCAGCTGCAGACATTAGAAAAAA-CTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
              ||||||| |||||||||||||||||||||| |||||||||||||| ||||||| ||||
Db        337 TCACTCATTCAGCTGCAGACATTAGAAAAAAGCTTCAAAAGTCTGCCTTAGGCCCAGAGC 278

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
              |||||||||||||||||||||||||||||:|| |||||||||||||||||||||| ||
Db        277 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCAGTTTTTTATAATAGAGATCAGGCGG 218

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
              ||||||| |||| |||||||| ||| |||| ||||||||||||||||||||| ||||||
Db        217 AGCAGGCAGAACGGGACAAACAGGA-TAAAGAAAAGGCCACCGCTTTAGTCATGGCCCTC 159

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 99

Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |
Db          98 AGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAAT 39

Qy       1922 AAGCCG-CCCCTTCGTCCATGCCCCTTATTTCAAGGGA 1958
              |||| | |||| || |||||||||||||| ||||||||
Db          38 AAGCTGCCCCCCTCATCCATGCCCCTTATGTCAAGGGA 1

RESULT 237
US-09-925-065A-948057/c
; Sequence 948057, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948057
;   LENGTH: 620
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948057

  Query Match             6.4%;  Score 481.5;  DB 1;  Length 620;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  548;  Conservative    3;  Mismatches   59;  Indels    7;  Gaps    3;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              |||||||||||||||||||| ||||||| ||||||||||||||||||| |||||||||
Db        613 AAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTG 554

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              |||||:|| |||||||||||||||||||||||||||||||| ||||   ||||||   |
Db        553 GCAACTTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGA 494

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              | |:|||||||||||| |||||||||| ||||||||||||||||||||||||||||| ||
Db        493 TTAMAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGG 434

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              |||||||||  |||| ||||| ||||||| ||||||||||||||||||||||||||||
Db        433 AAAAGGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACA 374

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||
Db        373 AGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTT 314

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || |||||||||||||||||||||||.|||| |||||||  |||||  | |||||||||
Db        313 ATGTCAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGA 254

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||
Db        253 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCA 194

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
              |||||||||||||||||| |||  |||||||||||||||||||||||||| | |   ||
Db        193 TGCCATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTG 134

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
              |||||||||||||||||| | ||||| |||||||||| ||||||| ||||||||||||
Db        133 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTC 74

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
              |||||||||||| ||| |     ||| | || ||| |||||||||||||  ||:||||
Db          73 CAGATCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCC 18
```

```
Qy       2245 AGCCACTAAGTTATGAA 2261
              ||||||||||||  | ||
Db         17 AGCCACTAAGTTGTAAA 1


RESULT 238
US-09-925-065A-948057/c
; Sequence 948057, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948057
;   LENGTH: 620
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948057

  Query Match             6.4%;  Score 481.5;  DB 1;  Length 620;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  548;  Conservative    3;  Mismatches   59;  Indels    7;  Gaps    3;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              ||||||||||||||||||||| ||||||| |||||||||||||||||| |||||||||
Db        613 AAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTGTTGAACTTG 554

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              |||||:|| |||||||||||||||||||||||||||||||| ||||   ||||||   |
Db        553 GCAACTTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAATGAGACAAAACTGGA 494

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              | |:||||||||||||| |||||||||| |||||||||||||||||||||||||||| ||
Db        493 TTAMAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCCGG 434

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              |||||||||  |||| ||||| ||||||| |||||||||||||||||||||||||||||
Db        433 AAAAGGGAAAGCCTGGACAAATCGAATGCCCAATAGGGCTTGCTTCCAGTGCGGTCTACA 374

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||| | ||||| |||||| |||| ||||||||||||||||
Db        373 AGGACACTTTAAAAAAGATTGTCCGAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTT 314

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || |||||||||||||||||||||||||||||| |||||||  |||||  | |||||||||
Db        313 ATGTCAAGGGAATCACTGGAAGGCCCACTGCCTCAGGGGATGAAGGTTCTCTGAGTCAGA 254

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||| ||||
Db        253 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCA 194

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
              |||||||||||||||| |||  |||||||||||||||||||||||||| | |   ||
Db        193 TGCCATCACCCTCACAGAACCCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTG 134

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
              |||||||||||||||| | ||||| |||||||||| ||||||| ||||||||||||||

Db        133 TCTCCTGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCCAGACAACTGTCCTC 74

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCC 2244
              |||||||||||| ||| |     |||  | || ||| ||||||||||||||  ||:||||
Db         73 CAGATCTGTCAATATCC---AAGGGGTCCTAGGACAGGCAGTCACTAGATA-GTTCTCCC 18

Qy       2245 AGCCACTAAGTTATGAA 2261
              ||||||||||||  | ||
Db         17 AGCCACTAAGTTGTAAA 1


RESULT 239
US-09-864-761-7607/c
; Sequence 7607, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7607
;   LENGTH: 549
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC022555.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
```

```
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
US-09-864-761-7607

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||.|||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  |||||||||||||||||:||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: ||||||||||.|||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:||||||:|||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1


RESULT 240
US-09-925-065A-739053/c
; Sequence 739053, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 739053
;   LENGTH: 630
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-739053

  Query Match             6.4%;  Score 480.9;  DB 1;  Length 630;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  570;  Conservative    8;  Mismatches   50;  Indels    9;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
             |||||||| || ||||||||||||||||||:||||||||||| ||:|||||||||||
Db        630 AGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG 571

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                | |      |    |||||||||||| |||||||||| ||||||||||||||||
Db        570 ----ACGCAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACTTGTGTGCTA 515

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
              :|||||||||||||||||:|||  ||: || :|||||| |||||||||||||| |||||
Db        514 GAAGGACTAAGGAAAACTAGGAA-TAAGCCTGTGAATTATTCAATGATGTCCACTATAAC 456

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||| ||||||
Db        455 ACAGGGAAAGGG-AGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAATGAGGA 397

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
              ||| | |||| |||||||||||||||  |||||:||||||||| ||||||  ||||||||
Db        396 AGCATACCTCCCTGTCACCTGACTCTATTGAAGRCCAACTAATGTTAAAGGATAAGTTTA 337

Qy       1623 TCACTCAGTCAGCTGCAGACATTAGAAAAAA-CTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
              ||||||| |||||||||||||||||||||||| |||||||||||||| ||||||| ||||
Db        336 TCACTCATTCAGCTGCAGACATTAGAAAAAAGCTTCAAAAGTCTGCCTTAGGCCCAGAGC 277

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
              |||||||||||||||||||||||||||||||||||:|| |||||||||||||||||| ||
Db        276 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCAGTTTTTTATAATAGAGATCAGGCGG 217

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
              ||||||| |||| |||||||| ||| |||| |||||||||||||||||||||| |||||
Db        216 AGCAGGCAGAACGGGACAAACAGGA-TAAAGAAAAGGCCACCGCTTTAGTCATGGCCCTC 158

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 98

Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |
Db         97 AGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAAT 38

Qy       1922 AAGCCG-CCCCTTCGTCCATGCCCCTTATTTCAAGGG 1957
              |||| | |||| || |||||||||||||| |||||||
Db          37 AAGCTGCCCCCCTCATCCATGCCCCTTATGTCAAGGG 1


RESULT 241
US-09-925-065A-739053/c
; Sequence 739053, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 739053
;   LENGTH: 630
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-739053

  Query Match             6.4%;  Score 480.9;  DB 1;  Length 630;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  570;  Conservative    8;  Mismatches   50;  Indels    9;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
              |||||||| || ||||||||||||||||||:||||||||||| ||:|||||||||||||
Db        630 AGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG 571

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                | |       |    |||||||||||| |||||||||| ||||||||||||||||
Db        570 ----ACGCAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACTTGTGTGCTA 515

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
              :|||||||||||||||||:|||  ||: || :|||||| |||||||||||||| |||||
Db        514 GAAGGACTAAGGAAAACTAGGAA-TAAGCCTGTGAATTATTCAATGATGTCCACTATAAC 456

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
              |||||| ||||| |||||||||||||||||||||||||||||||||||||||||| ||||||
Db        455 ACAGGGAAAGGG-AGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCAATGAGGA 397

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
              ||| | |||| ||||||||||||||  |||||:||||||||| ||||||  ||||||||
Db        396 AGCATACCTCCCTGTCACCTGACTCTATTGAAGRCCAACTAATGTTAAAGGATAAGTTTA 337

Qy       1623 TCACTCAGTCAGCTGCAGACATTAGAAAAAA-CTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
              ||||||| |||||||||||||||||||||| |||||||||||||| ||||||| ||||
Db        336 TCACTCATTCAGCTGCAGACATTAGAAAAAAGCTTCAAAAGTCTGCCTTAGGCCCAGAGC 277

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
              ||||||||||||||||||||||||||||||:|| |||||||||||||||||||||| ||
Db        276 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCAGTTTTTTATAATAGAGATCAGGCGG 217

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
              ||||||| |||| |||||||| ||| |||| |||||||||||||||||||||| |||||
Db        216 AGCAGGCAGAACGGGACAAACAGGA-TAAAGAAAAGGCCACCGCTTTAGTCATGGCCCTC 158

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 98

Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |
Db          97 AGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAAT 38

Qy       1922 AAGCCG-CCCCTTCGTCCATGCCCCTTATTTCAAGGG 1957
              |||| | |||| || |||||||||||||| |||||||
Db          37 AAGCTGCCCCCCTCATCCATGCCCCTTATGTCAAGGG 1

RESULT 242
US-10-301-480-203869/c
; Sequence 203869, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203869
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-203869

  Query Match             6.4%;  Score 480.8;  DB 1;  Length 604;
  Best Local Similarity   89.9%;  Pred. No. 0;
  Matches  545;  Conservative   10;  Mismatches   47;  Indels    4;  Gaps    4;

Qy       4661 CCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGG 4720
              |||:||| |||||||||||||||:|| |||:|||||||||||||| |:|:||:|||||
Db        604 CCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGAGAAGACTGG 545

Qy       4721 CAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTA 4779
              ||||| |||||||||||||| ||||||| |||| |||||||||||||||||||  || ||
Db        544 CAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTAGTTCGGCTA 486

Qy       4780 RATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTA 4839
              :||||||| || ||||||| |:||||||||| ||| |||| |||||||||||| |||| ||
Db        485 GATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCCCGAGAGATA 426

Qy       4840 ATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGT 4899
              ||||||  |||||||||||||||||||||||||| ||||| ||||||||||||||||
Db        425 ATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGCTTACAGAGT 367

Qy       4900 GACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATA 4959
              |||||| |||| |||||| |||| |||||||||||||||||||||||||| ||||| |||
Db        366 GACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATA 307

Qy       4960 CGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAAT 5019
              | |||||||||||| ||||||| |||||||| |||| |||||| || ||||||||||||
Db        306 CAATATCACTTACACCGCGCCTGGAGGCCACAATCCTTAGGGAAAGTTGAGAAAATGAAT 247

Qy       5020 GAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYT 5079
              |||| ||||||| |||||| |||||||| |||||||||||||||||||| |||||||:|
Db        246 GAAAAACTCAAATGACATCCAAAAAAGCTAACCCAGGAAACCCACCTCACGTGGCCTGCT 187

Qy       5080 CTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCA 5139
              :|||||||||||||||| |||||||| |||| ||||||||||||||||||||| ||
Db        186 STGTTGCCTATAGCCTTACTAAGAATCTGAAACTGTCCCCAAAAAGCAGGACTTAGCTCA 127

Qy       5140 TACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACA 5198
              |||||||||||||||| |||||||| |||||||||||||||||||||||  | |||||
Db        126 TACGAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTGCTTGACTGAGAGACA 67

Qy       5199 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 5258
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         66 GCCAACTTAGTTGCAGACATCACTTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 7

Qy       5259 CAAGGA 5264
              ||||||
Db          6 CAAGGA 1

RESULT 243
US-10-301-480-817278/c
; Sequence 817278, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
```

```
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817278
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-817278

  Query Match             6.4%;  Score 480.8;  DB 1;  Length 604;
  Best Local Similarity   89.9%;  Pred. No. 0;
  Matches  545;  Conservative   10;  Mismatches   47;  Indels    4;  Gaps    4;

Qy       4661 CCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGG 4720
              |||:||| |||||||||||||||:|| |||:|||||||||||||||| |:|:||:|||||
Db        604 CCTTATCACCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGAGAAGACTGG 545

Qy       4721 CAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTA 4779
              ||||| |||||||||||| ||||||| |||| |||||||||||||||||||  || ||
Db        544 CAACTGGATTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTAGTTCGGCTA 486

Qy       4780 RATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTA 4839
              :||||||| || ||||||| |:||||||||| ||| |||| |||||||||||| |||| ||
Db        485 GATACTTTTACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCCCGAGAGATA 426

Qy       4840 ATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGT 4899
              ||||||  |||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db        425 ATAAAGATACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGCTTACAGAGT 367

Qy       4900 GACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATA 4959
              |||||| |||| |||||| |||| |||||||||||||||||||||||| ||||| |||
Db        366 GACAATGGCCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATA 307

Qy       4960 CGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAAT 5019
              | |||||||||||| ||||||| |||||||| |||| |||||| || ||||||||||||
Db        306 CAATATCACTTACACCGCGCCTGGAGGCCACAATCCTTAGGGAAAGTTGAGAAAATGAAT 247

Qy       5020 GAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYT 5079
              |||| ||||||| |||||| |||||||| |||||||||||||||||||||| ||||||:|
Db        246 GAAAAACTCAAATGACATCCAAAAAAGCTAACCCAGGAAACCCACCTCACGTGGCCTGCT 187

Qy       5080 CTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCA 5139
              :|||||||||||||||  ||||||||| |||| ||||||||||||||||||||| ||
Db        186 STGTTGCCTATAGCCTTACTAAGAATCTGAAACTGTCCCCAAAAAGCAGGACTTAGCTCA 127

Qy       5140 TACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACA 5198
              |||||||||||||||||| |||||||| |||||||||||||||||||||||  | |||||
Db        126 TACGAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTGCTTGACTGAGAGACA 67

Qy       5199 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 5258
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          66 GCCAACTTAGTTGCAGACATCACTTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 7

Qy       5259 CAAGGA 5264
              ||||||
Db           6 CAAGGA 1


RESULT 244
US-10-301-480-382968
; Sequence 382968, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 382968
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-382968

  Query Match             6.4%;  Score 479.9;  DB 1;  Length 612;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  533;  Conservative   24;  Mismatches   55;  Indels    9;  Gaps    3;

Qy       2500 CCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCC 2559
              |||| |||||||||||||||||||||||||||||| ||||   ||||| |||| ||| ||
Db          1 CCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCA 60

Qy       2560 AAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCA 2619
              |||| |||||||||||| |||||||||||||| |||||| ||||||||  | |||||||
Db         61 AAAAAATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCA 120

Qy       2620 GTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCA 2679
              || | |       | |   |  | ||||||||||  |||||||||| ||||||||||
Db        121 ATACTCC-------AACTTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCA 172

Qy       2680 AGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTT 2739
              |||||||||||||||||| ||||| ||||||| | |||||||||||||||||| |||||
Db        173 AGATCTCAGGATTATCAAT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTT 231

Qy       2740 ATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGG 2799
              | | | ||::|||||||||||||||||||||||||||||||||:||||||||||:|||
Db        232 ACATTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGG 291

Qy       2800 ATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATA 2859
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATC 351

Qy       2860 CTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYC 2919
              ||||||||||:  |||||||||||||||||:|||||||||||||||||||||||||:|
Db        352 CTTCAAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCC 411

Qy       2920 CCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTC 2979
              |||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db        412 TCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTC 471

Qy       2980 TTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATC 3039
              |||||||||:|||: ||||||||||||||||:|   ||::|||||||||||||||   ||
Db        472 TTGTCCTTCAGTAYATGGATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGATGTC 531

Qy       3040 AAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSAR 3099
              ||||||| |||||:|||||:| |||||||  :|||| ||||||||:||||| |||||:|:
Db        532 AAGCCACACAAGCGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTTTCAAACCAA 591

Qy       3100 ARGCTCARCTCTGCTCACAGC 3120
              |:|||||:|||||||||||||
Db        592 AGGCTCAGCTCTGCTCACAGC 612


RESULT 245
US-10-301-480-996377
; Sequence 996377, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 996377
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-996377

  Query Match             6.4%;  Score 479.9;  DB 1;  Length 612;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  533;  Conservative   24;  Mismatches   55;  Indels    9;  Gaps    3;

Qy       2500 CCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCC 2559
             |||| |||||||||||||||||||||||||||||| ||||   ||||| |||| ||| ||
Db          1 CCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCA 60

Qy       2560 AAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCA 2619
             |||| |||||||||||| |||||||||||||| |||||| |||||||   |  |||||||
Db         61 AAAAAATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCA 120

Qy       2620 GTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCA 2679
              || | |     | | |   | ||||||||||  |||||||| ||||||||||
Db        121 ATACTCC-------AACTTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCA 172

Qy       2680 AGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTT 2739
             |||||||||||||||||| ||||| ||||||| | |||||||||||||||||| |||||
Db        173 AGATCTCAGGATTATCAAT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTT 231

Qy       2740 ATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGG 2799
             | | | ||::|||||||||||||||||||||||||||||||||||:|||||||||||:|||
Db        232 ACATTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGG 291

Qy       2800 ATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATA 2859
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATC 351

Qy       2860 CTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYC 2919
             ||||||||||:  |||||||||||||||||:|||||||||||||||||||||||||||:|
Db        352 CTTCAAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCC 411

Qy       2920 CCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTC 2979
             |||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db        412 TCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTC 471

Qy       2980 TTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATC 3039
             |||||||||:|||: |||||||||||||||||:|   ||::|||||||||||||||   ||
Db        472 TTGTCCTTCAGTAYATGGATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGATGTC 531

Qy       3040 AAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSAR 3099
             ||||||| |||||:|||||:| |||||||  :|||| ||||||||:|||| |||||:|:
Db        532 AAGCCACACAAGCGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTTTCAAACCAA 591

Qy       3100 ARGCTCARCTCTGCTCACAGC 3120
              |:||||||||||||||||
Db        592 AGGCTCAGCTCTGCTCACAGC 612


RESULT 246
US-09-925-065A-925326/c
; Sequence 925326, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925326
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925326

  Query Match             6.4%;  Score 479.9;  DB 1;  Length 644;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  530;  Conservative   33;  Mismatches   60;  Indels    9;  Gaps    3;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||| :|||:|  || ||||||||||||||||: |||||||:|||||||||||||||
Db        644 AATAGCTCTCACTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGC 585

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||: ||||||| | |||||||||::|||||||:|:|||||||||||||
Db        584 CTCTAAGTATGCTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATT 525

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||: ||:|||||||  |||:|||||||||||||| |||||||:||||| |:|| :
Db        524 CCTAACTTCCAAGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCY 465

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
             :|||||||||||||:||||||| :|||||||||||||: |||||||||      |||||:|
Db        464 TGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGA 410

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             ||||||| ||:|| :|||:|||||||  :||||||||| :|||||||  ||||||||||
Db        409 AAGGGAAGTAGAAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGA 351

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
             || |||||||||||||||||  ||      | |||||||||||||||||| || |||||||
Db        350 CCTTCCATTAGAAATGCTTATAGAAGGATGCCTAGTATAGGGTAATCCCCTCTGGGAAAC 291

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
             |||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| ||
Db        290 CAAGCCCCAGTACTCAGCAGGAAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCT 231

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             ||||||  || || |||||||||| ||||| ||||||||||  |||||| ||| ||||
Db        230 CCCCTCCAGATGGCTAGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACA 171

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             |||||||||||||||||||| |||||||| |||||||||| ||||||||||||||:||
Db        170 GAAATTACTTAAAACCCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGAT 111

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGC 4629
             |||||||| |||||||||||||||||||||||||||||||||||| |:|||:||||||
Db        110 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGG 51

Qy       4630 CTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||||:|||||||:|:|||||||||||||
Db         50 CTGTGAAGTGTGCCAAAGAAATAATCCCCTGC 19


RESULT 247
US-09-925-065A-925326/c
; Sequence 925326, Application US/09925065A·
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925326
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925326

  Query Match             6.4%;  Score 479.9;  DB 1;  Length 644;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  530;  Conservative   33;  Mismatches   60;  Indels    9;  Gaps    3;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             |||||| :|||:|  || |||||||||||||||||: |||||||:|||||||||||||||
Db        644 AATAGCTCTCACTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGC 585

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:||||||: ||||||| | |||||||||||::||||||||:|:||||||||||||
Db        584 CTCTAAGTATGCTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATT 525

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||: ||:|||||||  |||:|||||||||||| |||||||:|||| |:|| :
Db        524 CCTAACTTCCAAGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCY 465

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
             :|||||||||||||:|||||||| :|||||||||||||: |||||||||      |||||:|
Db        464 TGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGA 410

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             ||||||| ||:|| :|||:|||||||  :|||||||| :|||||||  ||||||||||
Db        409 AAGGGAAGTAGAAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGA 351

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
             || |||||||||||||||||  ||      | |||||||||||||||||| || |||||||
Db        350 CCTTCCATTAGAAATGCTTATAGAAGGATGCCTAGTATAGGGTAATCCCCTCTGGGAAAC 291

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
             |||||||||||||||||||||| ||| ||||| || |||||||| |||    | ||| ||
Db        290 CAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCT 231

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             ||||||  || || |||||||||| ||||| |||||||||  |||||| ||| ||||
Db        230 CCCCTCCAGATGGCTAGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACA 171

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             |||||||||||||||||||| |||||||| |||||||||| ||||||||||||||:||
Db        170 GAAATTACTTAAAACCCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGAT 111

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGC 4629
             |||||||| |||||||||||||||||||||||||||||||||||||| |:|||:||||||
Db        110 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGG 51

Qy       4630 CTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||||:|||||||:|:|||||||||||||||
Db          50 CTGTGAAGTGTGCCAAAGAAATAATCCCCTGC 19


RESULT 248
US-09-925-065A-936791/c
; Sequence 936791, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936791
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936791

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 659;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  593;  Conservative    3;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1656 TCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTC 1715
             ||||||||| ||| |||||| ||||||||| |||||||| |||||||||||||||||:||
Db        659 TCAAAAGTCCGCCTTAGGCCTGGAGCAAAATTTAGAAACACTATTGAACTTGGCAACCTC 600

Qy       1716 GGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAA 1775
              ||| |||||||||||||||||||||||||||||| |  |||||||||||| |||| |||
Db        599 AGTTGTTTATAATAGAGATCAGGAGGAGCAGGCGG-ATTGGACAAACGGGA-TAAACAAA 542

Qy       1776 AGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGA 1835
              |||||||  |||||||||  |||||||||||| ||||||||||||||||||||||||
Db        541 AGGCCACTTCTTTAGTCATAGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGGAAA----- 487

Qy       1836 AAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACT 1895
                 ||||||||  |||||||||||||||||||||||||||| | |||||||||  ||||
Db        486 ------GGGCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTGTGATCTACAAGGGCACT 433

Qy       1896 TTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAG 1955
              ||||| |||||||||||||||||| |||||  |||| | ||||||||||||||| |||||
Db        432 TTAAAGAAGATTGTCCAAGTAGAAATAAGCTACCCCCTTGTCCATGCCCCTTATGTCAAG 373

Qy       1956 GGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTA 2015
              |||||||||||||||| ||||||||||||  |||||| | ||||||||||||||||||
Db        372 GGAATCACTGGAAGGCCTACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTA 313

Qy       2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 2075
              |||||| ||||||||||:| |||||||||||||||||||||| |||| ||||||||||||
Db        312 ACCAGAAGATCCAGCAGYAAGACTGAGGGTGCCTGGGGCAAGTGCCAGCCCATGCCATCA 253

Qy       2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGG 2132
              |||||||||||||| |||||||  |||||||||||||||||||| | |   ||||||||||
Db        252 CCCTCACAGAGCCCCGGGTATGTGTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGG 193

Qy       2133 ACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTG 2192
              ||||||| |||| ||||| |||||||||||| |||| |||||||||||||||||||||||
Db        192 ACACTGGCGCGGCCTTCTCAGTCTTACTCTCCTGTTCCGGACAACTGTCCTCCAGATCTG 133

Qy       2193 TCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTA 2252
              ||||||| |    | |||   | || ||| |||||||||||||||| ||:|||||| ||||
Db        132 TCACTATCC---GAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTCTCCCAGCAACTA 77

Qy       2253 AGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCC 2312
              |||| || |||||||| ||| | |||||||||||||||||||||||||| |||||||||
Db         76 AGTTGTG-ACTGGGGAACTTCACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCC 18

Qy       2313 CACTACCTTGTTAGGGA 2329
              |||| ||||||||||||
Db         17 CACTCCCTTGTTAGGGA 1
```

```
RESULT 249
US-09-925-065A-936791/c
; Sequence 936791, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936791
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936791

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 659;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  593;  Conservative    3;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1656 TCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTC 1715
             ||||||||| ||| |||||| ||||||||| |||||||| ||||||||||||||||||:||
Db        659 TCAAAAGTCCGCCTTAGGCCTGGAGCAAAATTTAGAAACACTATTGAACTTGGCAACCTC 600

Qy       1716 GGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAA 1775
              ||| |||||||||||||||||||||||||||| |  ||||||||||| |||| |||
Db        599 AGTTGTTTATAATAGAGATCAGGAGGAGCAGGCGG-ATTGGACAAACGGGA-TAAACAAA 542

Qy       1776 AGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGA 1835
             |||||||  |||||||||  |||||||||||| ||||||||||||||||||||||
Db        541 AGGCCACTTCTTTAGTCATAGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGGAAA----- 487

Qy       1836 AAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACT 1895
                 ||||||||  |||||||||||||||||||||||||| | |||||||||| ||||
Db        486 ------GGGCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTGTGATCTACAAGGGCACT 433

Qy       1896 TTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAG 1955
              ||||| |||||||||||||||||| |||||  |||| | ||||||||||||||| |||||
Db        432 TTAAAGAAGATTGTCCAAGTAGAAATAAGCTACCCCCTTGTCCATGCCCCTTATGTCAAG 373

Qy       1956 GGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTA 2015
              |||||||||||||||| |||||||||||||  |||||| | ||||||||||||||||||
Db        372 GGAATCACTGGAAGGCCTACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTA 313

Qy       2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 2075
              |||||| ||||||||||:| |||||||||||||||||||||| |||| ||||||||||||
Db        312 ACCAGAAGATCCAGCAGYAAGACTGAGGGTGCCTGGGGCAAGTGCCAGCCCATGCCATCA 253

Qy       2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGG 2132
              |||||||||||||| |||||||  |||||||||||||||||||| | |   ||||||||||
Db        252 CCCTCACAGAGCCCCGGGTATGTGTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGG 193

Qy       2133 ACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTG 2192
              ||||||| |||| ||||| |||||||||| |||| ||||||||||||||||||||||||
Db        192 ACACTGGCGCGGCCTTCTCAGTCTTACTCTCCTGTTCCGGACAACTGTCCTCCAGATCTG 133

Qy       2193 TCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTA 2252
              ||||||| |   | |||  | || ||| |||||||||||||||| ||:|||||||| ||||

Db        132 TCACTATCC---GAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTCTCCCAGCAACTA 77

Qy       2253 AGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCC 2312
              |||| || |||||||| ||| | |||||||||||||||||||||||||| |||||||||
Db         76 AGTTGTG-ACTGGGGAACTTCACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCC 18

Qy       2313 CACTACCTTGTTAGGGA 2329
              |||| ||||||||||||
Db         17 CACTCCCTTGTTAGGGA 1


RESULT 250
US-09-925-065A-936792/c
; Sequence 936792, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936792
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936792

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 659;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  593;  Conservative    3;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1656 TCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTC 1715
             ||||||||| ||| |||||| ||||||||| |||||||| ||||||||||||||||||:||
Db        659 TCAAAAGTCCGCCTTAGGCCTGGAGCAAAATTTAGAAACACTATTGAACTTGGCAACCTC 600

Qy       1716 GGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAA 1775
              ||| |||||||||||||||||||||||||||| |  ||||||||||| |||| |||
Db        599 AGTTGTTTATAATAGAGATCAGGAGGAGCAGGCGG-ATTGGACAAACGGGA-TAAACAAA 542

Qy       1776 AGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGA 1835
             |||||||  |||||||||  |||||||||||| :|||||||||||||||||||||
Db        541 AGGCCACTTCTTTAGTCATAGCCCTCAGGCAAGCRGACTTTGGAGGCTCTGGAAA----- 487

Qy       1836 AAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACT 1895
                 ||||||||  |||||||||||||||||||||||||| | |||||||||| ||||
Db        486 ------GGGCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTGTGATCTACAAGGGCACT 433

Qy       1896 TTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAG 1955
              ||||| |||||||||||||||||| |||||  |||| | ||||||||||||||| |||||
Db        432 TTAAAGAAGATTGTCCAAGTAGAAATAAGCTACCCCCTTGTCCATGCCCCTTATGTCAAG 373

Qy       1956 GGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTA 2015
              |||||||||||||||| |||||||||||||  |||||| | ||||||||||||||||||
Db        372 GGAATCACTGGAAGGCCTACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTA 313

Qy       2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 2075
              |||||| |||||||||||| |||||||||||||||||||||| |||| ||||||||||||
Db        312 ACCAGAAGATCCAGCAGCAAGACTGAGGGTGCCTGGGGCAAGTGCCAGCCCATGCCATCA 253

Qy       2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGG 2132
```

```
                |||||||||||||||| |||||||  |||||||||||||||||||| | |   ||||||||||
Db        252 CCCTCACAGAGCCCCGGGTATGTGTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGG 193

Qy       2133 ACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTG 2192
                ||||||| |||| ||||| |||||||||| |||| ||||||||||||||||||||||||
Db        192 ACACTGGCGCGGCCTTCTCAGTCTTACTCTCCTGTTCCGGACAACTGTCCTCCAGATCTG 133

Qy       2193 TCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTA 2252
                ||||||| |   | |||  | || ||| |||||||||||||||| ||:||||||| ||||
Db        132 TCACTATCC---GAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTCTCCCAGCAACTA 77

Qy       2253 AGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCC 2312
                |||| || |||||||| ||| | ||||||||||||||||||||||||||| |||||||||
Db         76 AGTTGTG-ACTGGGGAACTTCACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCC 18

Qy       2313 CACTACCTTGTTAGGGA 2329
                |||| ||||||||||||
Db         17 CACTCCCTTGTTAGGGA 1


RESULT 251
US-09-925-065A-936792/c
; Sequence 936792, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936792
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936792

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 659;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  593;  Conservative    3;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1656 TCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTC 1715
                |||||||| ||| |||||| |||||||| |||||||| |||||||||||||||||:||
Db        659 TCAAAAGTCCGCCTTAGGCCTGGAGCAAAATTTAGAAACACTATTGAACTTGGCAACCTC 600

Qy       1716 GGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAA 1775
                ||| |||||||||||||||||||||||||||||| |  ||||||||||| |||| |||
Db        599 AGTTGTTTATAATAGAGATCAGGAGGAGCAGGCGG-ATTGGACAAACGGGA-TAAACAAA 542

Qy       1776 AGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGA 1835
                |||||||  |||||||||| |||||||||||| :||||||||||||||||||||
Db        541 AGGCCACTTCTTTAGTCATAGCCCTCAGGCAAGCRGACTTTGGAGGCTCTGGAAA----- 487

Qy       1836 AAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACT 1895
                      ||||||||  |||||||||||||||||||||||||| | |||||||||| ||||
Db        486 ------GGGCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTGTGATCTACAAGGGCACT 433

Qy       1896 TTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAG 1955
                |||||  |||||||||||||||||| |||||  |||| | |||||||||||||||| |||||
Db        432 TTAAAGAAGATTGTCCAAGTAGAAATAAGCTACCCCCTTGTCCATGCCCCTTATGTCAAG 373

Qy       1956 GGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTA 2015
                |||||||||||||||| |||||||||||||  |||||| | ||||||||||||||||||
Db        372 GGAATCACTGGAAGGCCTACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAGCCACTA 313

Qy       2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 2075
                |||||| |||||||||||| |||||||||||||||||||||| |||| ||||||||||||
Db        312 ACCAGAAGATCCAGCAGCAAGACTGAGGGTGCCTGGGGCAAGTGCCAGCCCATGCCATCA 253

Qy       2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGG 2132
                |||||||||||||| |||||||  |||||||||||||||||||| | |   ||||||||||
Db        252 CCCTCACAGAGCCCCGGGTATGTGTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGG 193

Qy       2133 ACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTG 2192
                ||||||| |||| ||||| |||||||||| |||| ||||||||||||||||||||||||
Db        192 ACACTGGCGCGGCCTTCTCAGTCTTACTCTCCTGTTCCGGACAACTGTCCTCCAGATCTG 133

Qy       2193 TCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTA 2252
                ||||||| |   | |||  | || ||| |||||||||||||||| ||:||||||| ||||
Db        132 TCACTATCC---GAGGGGTCCTAGGACAGGCAGTCACTAGATAC-TTCTCCCAGCAACTA 77

Qy       2253 AGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCC 2312
                |||| || |||||||| ||| | ||||||||||||||||||||||||||| |||||||||
Db         76 AGTTGTG-ACTGGGGAACTTCACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCC 18

Qy       2313 CACTACCTTGTTAGGGA 2329
                |||| ||||||||||||
Db         17 CACTCCCTTGTTAGGGA 1


RESULT 252
US-10-029-386-8924/c
; Sequence 8924, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8924
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL137799.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P03357, EVALUE 2.00e-37
;   OTHER INFORMATION: EST_HUMAN HIT: AW967582.1, EVALUE 0.00e+00
US-10-029-386-8924

  Query Match             6.4%;  Score 477.3;  DB 1;  Length 559;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  489;  Conservative   33;  Mismatches   37;  Indels    1;  Gaps    1;

Qy       3840 TTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCT 3899
                | |||||||||||||| |||||| |||||:||||:|||||:|||||||||||:||||
Db        559 TGCTCAAACCTATGCCGCTCGAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGACC- 501

Qy       3900 TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATG 3959
                |||||||||||| |||||||||||||||||||||||||| ||||||||||||: |||||||
Db        500 TCAACTTGTATATTGATGGAAGTTCCTTTGTAGAAAAAAGACTTCGAAAAGCAGGGTATG 441

Qy       3960 CAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTY 4019
```

```
                |||||||||||||||||||||||:|||||| ||||||||||||||||||||||||| |||:
Db        440 CAGTGGTCAGTGATAATGGAATACTTGAAAATAATCCCCTCACTCCAGGAACTAGCGCTC 381

Qy       4020 AGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA 4079
              ||||:||| |||||||||||:| |:| |||||||| |||||| | |||:|||||||:||
Db        380 AGCTGGCAAAACTAATAGCCCTTACTCGGGCACTAAAATTAGAAAAAGAAAAAAGGGTAA 321

Qy       4080 ATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSA 4139
              | |||||||||||||||:|:||||||:|||||||| ||||||||||||::|||||||:|
Db        320 ACATATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGA 261

Qy       4140 RAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGA 4199
              :|||||||||||| |||||||||:|||:||||| ||||||:|||||||||||||| ||||
Db        260 GAGAAAGGGAATACCTAACTTCTGAGGGAACATCTATCAAACATCAGGAAGCCATCAGGA 201

Qy       4200 RATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATC 4259
              :|||||||: |||:|||||||||||||:||||||||:|||||||||||||:|||||||||
Db        200 GATTATTATCGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATC 141

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
              | |||||||||:||||||||||:|||:|||: ||||||| ||||||||||:|||||||
Db        140 AGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCGAATATTGAAGCCAAAAGAG 81

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
              |  |||||||||||||||||| |||||||||  ||   ||| |||||| ||||||||
Db          80 CCACAAGGCAGGACCCTCCATTAAAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATC 21

Qy       4380 CCTTCCGGGAAACCAAGCCC 4399
              || || ||||||||||||||
Db          20 CCCTCTGGGAAACCAAGCCC 1


RESULT 253
US-09-925-065A-769980
; Sequence 769980, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769980
;   LENGTH: 656
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-769980

  Query Match             6.4%;  Score 476.2;  DB 1;  Length 656;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  590;  Conservative   12;  Mismatches   52;  Indels   22;  Gaps    8;

Qy        790 CAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATT 849
              |||||||||||||||||||| || |||||||||||||||||||||||||||||| |||||
Db          1 CAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTCGCAATT 60

Qy        850 ATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAG 909
              |           | | |||||||||||||||||| ||||||||||||||||||||||||
Db         61 A-----------CGGTGTATGCCCTACAGGAAGCCCTCAGAGTCTACCTCCCTATCCCAG 109

Qy        910 CAT--CCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAA 967
              |||  |||||||||||||||::|||||||||||||||||||||||||||||| |||||||||
Db        110 CATCCCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAACGGTCCAAAA 169

Qy        968 GGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACC 1027
              |||||||||||||| ||||||||| ||||| |||||||||||| |||||||| |||||| ||
Db        170 GGAGATAGACAAATGGGTAAACAATGAACTAAAGAGTGCCAGTATTCCCCGATTATG-CC 228

Qy       1028 CCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTC 1087
              ||| |||||||||||||||| ||||| | ||||||||||||||||||||| |:||||::||
Db        229 CCTTCCAAGCAGTGGGAGGAGGAGAAATTGGCCCAGCCAGAGTGCATGTACCTTTTTCTC 288

Qy       1088 TCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCT 1147
              || |||||||||||||||| |||| |||| ||||||||||||||||||:||||||||||
Db        289 TCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACCCTGATGGCT 348

Qy       1148 ATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATG 1207
              ||||||:|||||||||||| |||| |||| |||| ||||||||||||||||||| |||
Db        349 ATATTGATGTTTTACAAGGGCTAGGGCAATCCTTTAATCTGACATGGAGAGATATA-ATG 407

Qy       1208 TCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGA 1267
              | ||||||| ||||||| |||||||||||||| ||||||||| :|||||||||||| ||
Db        408 TTACTGCTAGATCAGACGCTAACCCCAAATGAAAGAAGTGCCGYCATAACTGCAGCCCGA 467

Qy       1268 GRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAA 1327
              |:|||||||||||||||||||||||||||||||||||||||     |||||||||||| |||||
Db        468 GAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-AGGAA 524

Qy       1328 AGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACAC 1387
              ||| || ||||||||||||||||||:|||||||||||| ||:||||||||| ||   ||
Db        525 AGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTAGG---AC 581

Qy       1388 AGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGG 1447
              | | |     ||||| |||||||||||  | |||||||||||||| |||||| ||||
Db        582 ACAGAATCAGAACAT-GGAGATTGGTGCCACTGACATTTGCTAACTTGAGTGCTAAAAGG 640

Qy       1448 ACTAAGGAAAACTASG 1463
              ||||| ||||||||:|
Db        641 ACTAAAGAAAACTAGG 656


RESULT 254
US-09-925-065A-769980
; Sequence 769980, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769980
;   LENGTH: 656
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-769980

  Query Match             6.4%;  Score 476.2;  DB 1;  Length 656;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  590;  Conservative   12;  Mismatches   52;  Indels   22;  Gaps    8;
```

```
Qy        790 CAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATT 849
              ||||||||||||||||||||| || |||||||||||||||||||||||||||||| |||||
Db          1 CAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTCGCAATT 60

Qy        850 ATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAG 909
              |           | | |||||||||||||||||| ||||||||||||||||||||||||
Db         61 A-----------CGGTGTATGCCCTACAGGAAGCCCTCAGAGTCTACCTCCCTATCCCAG 109

Qy        910 CAT--CCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAA 967
              |||  ||||||||||||||||::|||||||||||||||||||||||||| |||||||||
Db        110 CATCCCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAACGGTCCAAAA 169

Qy        968 GGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACC 1027
              |||||||||||||| ||||||||| ||||| |||||||||| |||||||| |||||| ||
Db        170 GGAGATAGACAAATGGGTAAACAATGAACTAAAGAGTGCCAGTATTCCCCGATTATG-CC 228

Qy       1028 CCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTC 1087
              ||| ||||||||||||||| ||||| | |||||||||||||||||| |:||||::||
Db        229 CCTTCCAAGCAGTGGGAGGAGGAGAAATTGGCCCAGCCAGAGTGCATGTACCTTTTTCTC 288

Qy       1088 TCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCT 1147
              || |||||||||||||| |||| |||| |||||||||||||||||:||||||||||
Db        289 TCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACCCTGATGGCT 348

Qy       1148 ATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATG 1207
              ||||||:|||||||||||||| |||| |||| |||| |||||||||||||||||||| |||
Db        349 ATATTGATGTTTTACAAGGGCTAGGGCAATCCTTTAATCTGACATGGAGAGATATA-ATG 407

Qy       1208 TCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGA 1267
              | ||||||| ||||||| |||||||||||||| ||||||||| :|||||||||||| ||
Db        408 TTACTGCTAGATCAGACGCTAACCCCAAATGAAAGAAGTGCCGYCATAACTGCAGCCCGA 467

Qy       1268 GRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAA 1327
              |:|||||||||||||||||||||||||||||||||||||     |||||||||||| |||||
Db        468 GAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-AGGAA 524

Qy       1328 AGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACAC 1387
              ||| || ||||||||||||||||||:|||||||||||| ||:|||||||| ||    ||
Db        525 AGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTAGG---AC 581

Qy       1388 AGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGG 1447
              | | |     ||||| |||||||||||  | |||||||||||||||| |||||| ||||
Db        582 ACAGAATCAGAACAT-GGAGATTGGTGCCACTGACATTTGCTAACTTGAGTGCTAAAAGG 640

Qy       1448 ACTAAGGAAAACTASG 1463
              ||||| ||||||||:|
Db        641 ACTAAAGAAAACTAGG 656
```

RESULT 255
US-09-925-065A-785421/c
; Sequence 785421, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 785421
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-785421

Query Match 6.3%; Score 475.2; DB 1; Length 586;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 540; Conservative 5; Mismatches 38; Indels 10; Gaps 4;

```
Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
              ||:||||||||||||||||||:|||||    |||:|||||| |||||||||||||| |||
Db         586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
              |||| |||||| |||||||||||||| |||||||||||||||||||||||||||| |||||
Db         530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
              ||||| | ||||||  |||||||| |||  ||||||||||||||||||||||  ||| ||
Db         471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
              ||||||||| ||||||||||||||||   |||||||||||||||||||||| ||||||||
Db         413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
              |||||||||||||||||||||||||||||||| |||||||:| |||||||||||| ||||||||
Db         353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
              ||||||||||:|||||||||||| ||||| |||||||||| |||||||||||||||| |
Db         293 GAGGAGCAGGYGGAACAGGACAAATGGGATAAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
              |||||||||| |||||| |||||||||||||||||||||||||| ||||||||||||
Db         233 CCTCAGGCAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
              ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db         173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
              |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db         113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              ||||| ||||| |||||| | |||||||||||||||||||||||||||||| |
Db          53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1
```

RESULT 256
US-09-925-065A-785421/c
; Sequence 785421, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785421
;   LENGTH: 586
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785421

  Query Match             6.3%;  Score 475.2;  DB 1;  Length 586;
  Best Local Similarity   91.1%;  Pred. No. 0;
  Matches  540;  Conservative    5;  Mismatches   38;  Indels   10;  Gaps    4;

Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
             ||:||||||||||||||||||:|||||    |||:|||||| |||||||||||||| |||
Db        586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
             |||| |||||| |||||||||||||| |||||||||||||||||||||||||||| |||||
Db        530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
             ||||| | ||||||  |||||||| |||  |||||||||||||||||||||  ||| ||
Db        471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
             ||||||||| ||||||||||||||||   |||||||||||||||||||||| ||||||||
Db        413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
             |||||||||||||||||||||||||||||||| |||||||:| |||||||||||| |||||||||
Db        353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
             ||||||||||:|||||||||||||| ||||| |||||||||| |||||||||||||||| |
Db        293 GAGGAGCAGGYGGAACAGGACAAATGGGATAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
             |||||||||||| |||||| |||||||||||||||||||||||||||| ||||||||||||
Db        233 CCTCAGGCAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
             ||||||| ||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db        173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
             |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db        113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             ||||| ||||| |||||| | |||||||||||||||||||||||||||||| |
Db          53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1

RESULT 257
US-09-925-065A-103629/c
; Sequence 103629, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09

;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103629
;   LENGTH: 597
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-103629

  Query Match             6.3%;  Score 474.8;  DB 1;  Length 597;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  539;  Conservative    9;  Mismatches   46;  Indels    4;  Gaps    4;

Qy       4669 CCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GA 4727
             |||||||||||||||||:|| |||:|||||||||||||| |:|:||:|||||||||| ||
Db        596 CCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGAGAAGACTGGCAACTGGA 537

Qy       4728 TTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTT 4787
             |||||||||| ||||||| |||| |||||||||||||||||||  || ||:|||||||
Db        536 TTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTAGTTCGGCTAGATACTTT 478

Qy       4788 CACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGC 4847
              || ||||||| |:||||||||| ||| |||| |||||||||| |||| ||||||||
Db        477 TACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCCCGAGAGATAATAAAGAT 418

Qy       4848 ACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAG 4907
             |||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||| |
Db        417 ACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGCTTACAGAGTGACAATGG 359

Qy       4908 CCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCA 4967
             ||| |||||| |||| |||||||||||||||||||||||||| ||||| |||| ||||||
Db        358 CCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCA 299

Qy       4968 CTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACT 5027
             ||||||| ||||||| |||||||| |||| |||||| || |||||||||||||||| |||
Db        298 CTTACACCGCGCCTGGAGGCCACAATCCTTAGGGAAAGTTGAGAAAATGAATGAAAAACT 239

Qy       5028 CAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCC 5087
             |||| |||||| |||||||| |||||||||||||||||||||| |||||||:|:|||||||
Db        238 CAAATGACATCCAAAAAAGCTAACCCAGGAAACCCACCTCACGTGGCCTGCTSTGTTGCC 179

Qy       5088 TATAGCCTTAAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAAT 5147
             |||||||||| |||||||| |||| |||||||||||||||||||||| ||||||||||
Db        178 TATAGCCTTACTAAGAATCTGAAACTGTCCCCAAAAAGCAGGACTTAGCTCATACGAAAT 119

Qy       5148 GCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTT 5206
             |||||||||| |||||||| |||||||||||||||||||||| | ||||||||||||
Db        118 GCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTGCTTGACTGAGAGACAGCCAACTT 59

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGA 5264
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db          58 AGTTGCAGACATCACTTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGA 1

RESULT 258
US-09-925-065A-103629/c
; Sequence 103629, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103629
;   LENGTH: 597
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-103629

  Query Match             6.3%;  Score 474.8;  DB 1;  Length 597;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  539;  Conservative    9;  Mismatches   46;  Indels    4;  Gaps    4;

Qy       4669 CCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GA 4727
             ||||||||||||||||:|| |||:|||||||||||||||| |:|:||:|||||||||| ||
Db        596 CCAAGCTCCTTCAGGAGAAAAAAGAACAGGCCATTACCCAGGAGAAGACTGGCAACTGGA 537

Qy       4728 TTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTT 4787
             |||||||||| ||||||| |||| |||||||||||||||||||  || ||:|||||||
Db        536 TTTTACCCACATGCCCAAATCTCA-GGATTTCAGTATCTACTAGTTCGGCTAGATACTTT 478

Qy       4788 CACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGC 4847
               || ||||||| |:||||||||| ||| |||| |||||||||||| |||| |||||||||
Db        477 TACTGGTTGGGGAGAGGCCTTCCACTGCAGGAAAGAAAAGGCCCGAGAGATAATAAAGAT 418

Qy       4848 ACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAG 4907
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||||| |
Db        417 ACTAGTTCATGAAATAATTCCCAGATTCAGACTT-CCCGAGGCTTACAGAGTGACAATGG 359

Qy       4908 CCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCA 4967
             ||| |||||| |||| |||||||||||||||||||||||| ||||| |||| ||||||
Db        358 CCCCGCTTTCAAGGCGACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCA 299

Qy       4968 CTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACT 5027
             ||||||| ||||||| |||||||| |||| |||||| || ||||||||||||||| |||
Db        298 CTTACACCGCGCCTGGAGGCCACAATCCTTAGGGAAAGTTGAGAAAATGAATGAAAAACT 239

Qy       5028 CAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCC 5087
             |||| |||||| |||||||| |||||||||||||||||||||| |||||||:|:|||||||
Db        238 CAAATGACATCCAAAAAAGCTAACCCAGGAAACCCACCTCACGTGGCCTGCTSTGTTGCC 179

Qy       5088 TATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAAT 5147
             |||||||||| |||||||| |||| |||||||||||||||||||||| ||||||||||
Db        178 TATAGCCTTACTAAGAATCTGAAACTGTCCCCAAAAAGCAGGACTTAGCTCATACGAAAT 119

Qy       5148 GCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTT 5206
             |||||||||| |||||||| |||||||||||||||||||||||| | ||||||||||||
Db        118 GCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTGCTTGACTGAGAGACAGCCAACTT 59

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGA 5264
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          58 AGTTGCAGACATCACTTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGA 1


RESULT 259
US-09-925-065A-764810/c
; Sequence 764810, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 764810
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-764810

  Query Match             6.3%;  Score 474;  DB 1;  Length 614;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  546;  Conservative    8;  Mismatches   58;  Indels    4;  Gaps    4;

Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
             | |:|:||:|||||||||| |||||||||||||| ||||||| ||||||||||||||||||
Db        614 CAGGAGAAGACTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 555

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
             |||||||||||||||:|||||||||| |||||||| :|||||||||| ||||||||||||
Db        554 TACTAGTCTGGGTAGATACTTTCACTGGTTGGGCGGAGGCCTTCCCTTGTAGGACAGAAC 495

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
             ||||||| |||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db        494 AGGCCCATGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTTGGACTTCCCC 435

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
             |||||||||||||| ||  ||||  ||||| |||| ||||||||||||||||||||||||
Db        434 AAGGCTTACAGAGTGATAACGGCCCCACTTTCAAGGCTACAGTAACCCAGGGAGTATCCC 375

Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
             || | ||||  |||| |||||||||||||| ||| | |||||||| |||||| |||| |
Db        374 AGACATTAGACATACAATATCACTTACACTGAGCCCGGAGGCCACAATCCTCA-GGAAAG 316

Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACC 5065
             | |||||||||||||||: |||||| |||||||||||||| |||| | ||||||||||
Db        315 TTGAGAAAATGAATGAAACGCTCAAATGACATCTAAAAAAGCTAACCTAAGAAACCCACC 256

Qy       5066 TCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAG 5125
             || |||||  ||:|||||||||||||||||  |||||||| | |||| ||||||||||
Db        255 TCKCATGGTTTGCTCTGTTGCCTATAGCCTTAGTAAGAATCCGAAACTCTCCCCAAAAAG 196

Qy       5126 CAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGC 5185
             | ||||| |||||||||||||||||||||| |||||||| ||||||||||||||||||
Db        195 CGGGACTCAGCCCATACGAAATGCTGTATGGACGGCCCTTCCTAACCAATGACCTTGTGC 136

Qy       5186 TTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAG 5244
             |||||| | |||  |||||||||||||||| ||| ||||||||||||||||||||||||
Db        135 TTGACCTAGAGATGGCCAACTTAGTTGCAGATATC-CCTCCTTAGCCAAATATCAACAAG 77

Qy       5245 TTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWW 5304
             |||||||||||| ||| ||||||| ||||||||| |||||||| ||| ||||||| ||:
Db         76 TTCTTAAAACATCACAGGGAACCTGTCCCTGAGAGGAGGGAAAGGAATTATTCCAACCTG 17

Qy       5305 GTGACATGGTATTAGT 5320
             ||||||||||||||||
Db         16 GTGACATGGTATTAGT 1


RESULT 260
US-09-925-065A-764810/c
; Sequence 764810, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
```

```
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 764810
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-764810

  Query Match             6.3%;  Score 474;  DB 1;  Length 614;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  546;  Conservative    8;  Mismatches   58;  Indels    4;  Gaps    4;

Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
              | |:|:||:|||||||||| |||||||||||| ||||||| ||||||||||||||||||
Db        614 CAGGAGAAGACTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 555

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
              |||||||||||||:|||||||||| |||||||| :|||||||||| ||||||||||||
Db        554 TACTAGTCTGGGTAGATACTTTCACTGGTTGGGCGGAGGCCTTCCCTTGTAGGACAGAAC 495

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
              ||||||| |||||||||||||||| |||||||||||||||||||||| ||||||||||
Db        494 AGGCCCATGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTTGGACTTCCCC 435

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
              |||||||||||||||| ||  ||||  ||||| |||| ||||||||||||||||||||||
Db        434 AAGGCTTACAGAGTGATAACGGCCCCACTTTCAAGGCTACAGTAACCCAGGGAGTATCCC 375

Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
              || | ||||  |||| |||||||||||||||| ||| | |||||||| |||||| |||| |
Db        374 AGACATTAGACATACAATATCACTTACACTGAGCCCGGAGGCCACAATCCTCA-GGAAAG 316

Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACC 5065
              | |||||||||||||||: |||||| |||||||||||||||| |||| | ||||||||||
Db        315 TTGAGAAAATGAATGAAACGCTCAAATGACATCTAAAAAAGCTAACCTAAGAAACCCACC 256

Qy       5066 TCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAG 5125
              || |||||  ||:|||||||||||||||||||  ||||||| | |||| ||||||||||
Db        255 TCKCATGGTTTGCTCTGTTGCCTATAGCCTTAGTAAGAATCCGAAACTCTCCCCAAAAAG 196

Qy       5126 CAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGC 5185
              | ||||| |||||||||||||||||||||||| |||||||| ||||||||||||||||||
Db        195 CGGGACTCAGCCCATACGAAATGCTGTATGGACGGCCCTTCCTAACCAATGACCTTGTGC 136

Qy       5186 TTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAG 5244
              |||||| | |||  |||||||||||||||| ||| ||||||||||||||||||||||||
Db        135 TTGACCTAGAGATGGCCAACTTAGTTGCAGATATC-CCTCCTTAGCCAAATATCAACAAG 77

Qy       5245 TTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWW 5304
              |||||||||||| ||| ||||||| ||||||||| |||||||| ||| ||||||| ||:
Db          76 TTCTTAAAACATCACAGGGAACCTGTCCCTGAGAGGAGGGAAAGGAATTATTCCAACCTG 17

Qy       5305 GTGACATGGTATTAGT 5320
              ||||||||||||||||
Db          16 GTGACATGGTATTAGT 1


RESULT 261
US-10-029-386-563/c
; Sequence 563, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 563
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL137799.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.32
;   OTHER INFORMATION: NT HIT: AL163280.2, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AU121943.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P26808, EVALUE 5.00e-41
US-10-029-386-563

  Query Match             6.3%;  Score 473.9;  DB 1;  Length 568;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  499;  Conservative   30;  Mismatches   36;  Indels    3;  Gaps    2;

Qy       2644 GCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGG 2703
              | |||||||||||  || ||||||| || |||||||||||||||||||||| |||| |
Db        565 GTACAGAAACCCAACGGGCAGTGGA-GGCTAGTGCAAGATCTCAGGATTATAAATGAGGT 507

Qy       2704 CCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAG 2763
              |||||| | ||||| |||||||||||| |||||||||| ||:::|||||||||||||||
Db        506 A--TTGTCCCTCTATACTCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATACCAG 449

Qy       2764 AGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTAC 2823
              |||||||||||||||||||:||||||||||:|||||||||| |||||||||| ||||||
Db        448 AGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCTCTGTAC 389

Qy       2824 ATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCA 2883
              | ||| |||||||||||||||||||||||||||||| |||| |||||:| ||||||||||
Db        388 AACCTAACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCGAACCCAACGTCTCAACTCA 329

Qy       2884 CCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAG 2943
              ||||||||:|||||||||||||||| |||||||: |||||||||||||||||| |||
Db        328 CCTGGACTGTTTTACCCCAAGGGTTCCGGGATAGCTCCCATCTATTTGGCCAGGCGTTAA 269

Qy       2944 CCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATT 3003
              |||||||||||||:||:|::|||||||| |||||||||||||||:|||:||||||||||
Db        268 CCCAAGACTTGAGCCAGTTCTCATACCTTGACACTCTTGTCCTTCAGTATGTGGATGATT 209

Qy       3004 TACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATT 3063
              | |||||:||: ||::|||||||||||||||||||||||| ||||||||:||||| ||||
Db        208 TGCTTTTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCTACCCAAGCGCTCTTGAATT 149

Qy       3064 TCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGG 3123
              |||||||:|||||||||||||:|||||||||| |:|:|||||:|||||||||| ||||
Db        148 TCCTCGCCACCTGTGGCTACAAGGTTTCCAAACAAAAGGCTCAGCTCTGCTCACAACAGG 89

Qy       3124 TTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCC 3183
              |||||| |||||||:||||:|||||||||||||||||:|||||||||| |||||:::||||
Db         88 TTAAATGCTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTAAGGAATGTATCC 29

Qy       3184 AGCCTATACTGGCTTATCCTCATCYCAA 3211
              |||| ||||||||||||||||||:|||
Db         28 AGCCCATACTGGCTTATCCTCATCCCAA 1


RESULT 262
US-09-925-065A-785419/c
; Sequence 785419, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
```

```
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785419
;   LENGTH: 586
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785419

  Query Match             6.3%;  Score 473.6;  DB 1;  Length 586;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  539;  Conservative    5;  Mismatches   39;  Indels   10;  Gaps    4;

Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
             ||:||||||||||||||||||:|||||    |||:|||||| ||||||||||||||| |||
Db        586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
             |||| |||||| |||||||||||||| |||||||||||||||||||||||||||||| |||||
Db        530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
             ||||| | |||||||  |||||||| |||  |||||||||||||||||||||||  ||| ||
Db        471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
             |||||||| ||||||||||||||||    |||||||||||||||||||||| ||||||||
Db        413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
             |||||||||||||||||||||||||||||| |||||||:| |||||||||||| |||||||||
Db        353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
             |||||||||| |||||||||||| ||||| |||||||||| |||||||||||||||| |
Db        293 GAGGAGCAGGTGGAACAGGACAAATGGGATAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
             |||||||:||| |||||| ||||||||||||||||||||||||| ||||||||||||
Db        233 CCTCAGGMAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
             ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db        173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
             |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db        113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             ||||| |||||| |||||| | ||||||||||||||||||||||||||||||| |
Db         53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1

RESULT 263
US-09-925-065A-785419/c
; Sequence 785419, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.

;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785419
;   LENGTH: 586
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785419

  Query Match             6.3%;  Score 473.6;  DB 1;  Length 586;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  539;  Conservative    5;  Mismatches   39;  Indels   10;  Gaps    4;

Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
             ||:||||||||||||||||||:|||||    |||:|||||| ||||||||||||||| |||
Db        586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
             |||| |||||| |||||||||||||| |||||||||||||||||||||||||||||| |||||
Db        530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
             ||||| | |||||||  |||||||| |||  |||||||||||||||||||||||  ||| ||
Db        471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
             |||||||| ||||||||||||||||    |||||||||||||||||||||| ||||||||
Db        413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
             |||||||||||||||||||||||||||||| |||||||:| |||||||||||| |||||||||
Db        353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
             |||||||||| |||||||||||| ||||| |||||||||| |||||||||||||||| |
Db        293 GAGGAGCAGGTGGAACAGGACAAATGGGATAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
             |||||||:||| |||||| ||||||||||||||||||||||||| ||||||||||||
Db        233 CCTCAGGMAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
             ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db        173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
             |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db        113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
             ||||| |||||| |||||| | ||||||||||||||||||||||||||||||| |
Db         53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1

RESULT 264
US-09-925-065A-785420/c
; Sequence 785420, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785420
;   LENGTH: 586
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785420

  Query Match             6.3%;  Score 473.6;  DB 1;  Length 586;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  539;  Conservative    5;  Mismatches   39;  Indels   10;  Gaps    4;

Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
                ||:||||||||||||||||||:|||||    |||:|||||| |||||||||||||| |||
Db        586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
                |||| |||||| |||||||||||||| |||||||||||||||||||||||||| |||||
Db        530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
                |||||  | ||||||  |||||||| |||  ||||||||||||||||||||  ||| ||
Db        471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
                |||||||| ||||||||||||||||   |||||||||||||||||||| ||||||||
Db        413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
                |||||||||||||||||||||||||| |||||||:| |||||||||||| ||||||||
Db        353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
                |||||||||| |||||:||||||| ||||| |||||||||| |||||||||||||| |
Db        293 GAGGAGCAGGTGGAACRGGACAAATGGGATAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
                |||||||||| |||||| |||||||||||||||||||||||||| ||||||||||||
Db        233 CCTCAGGCAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
                ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db        173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
                |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db        113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                ||||| ||||| |||||| | |||||||||||||||||||||||||||||| |
Db          53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1

RESULT 265
US-09-925-065A-785420/c
; Sequence 785420, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785420
;   LENGTH: 586
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785420

  Query Match             6.3%;  Score 473.6;  DB 1;  Length 586;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  539;  Conservative    5;  Mismatches   39;  Indels   10;  Gaps    4;

Qy       1441 TASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATA 1500
                ||:||||||||||||||||||:|||||    |||:|||||| |||||||||||||| |||
Db        586 TAGAAGGACTAAGGAAAACTAGGAAGA----CTATGAATTATTCAATGATGTCCACTATA 531

Qy       1501 ACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560
                |||| |||||| |||||||||||||| |||||||||||||||||||||||||| |||||
Db        530 ACAC-GGGGAAAGGAAGAAAATCCTACCGCCTTTCTGGAGAGACTAAGGGAGGCGTTGAG 472

Qy       1561 GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620
                |||||  | ||||||  |||||||| |||  ||||||||||||||||||||  ||| ||
Db        471 GAAGCATACCTCTC--TCACCTGATTCTATTGAAGGCCAACTAATCTTAAAGGATAAATT 414

Qy       1621 TATCACTCAGTCAGCTGCAGACATTAG---AAAAAACTTCAAAAGTCTGCCGTAGGCCCG 1677
                |||||||| ||||||||||||||||   |||||||||||||||||||| ||||||||
Db        413 TATCACTCAATCAGCTGCAGACATTAGAAAAAAAAAACTTCAAAAGTCTGCCATAGGCCCG 354

Qy       1678 GAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAG 1737
                |||||||||||||||||||||||||| |||||||:| |||||||||||| ||||||||
Db        353 GAGCAAAACTTAGAAACCCTATTGAACGTGGCAACCTTGGTTTTTTATAAAAGAGATCAG 294

Qy       1738 GAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGAC 1797
                |||||||||| |||||:||||||| ||||| |||||||||| |||||||||||||| |
Db        293 GAGGAGCAGGTGGAACRGGACAAATGGGATAAAAAAAAAGGCTACCGCTTTAGTCATGGC 234

Qy       1798 CCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCC 1857
                |||||||||| |||||| |||||||||||||||||||||||||| ||||||||||||
Db        233 CCTCAGGCAAGCGGACTTCGGAGGCTCTGGAAAAGGGAAAAGCTGGGAAAATTGAATGCC 174

Qy       1858 TAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAG 1917
                ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db        173 TAATAGGCCTTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAAAAGATTGTCCAAGTAG 114

Qy       1918 AAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTG 1977
                |||||||| ||||| | |||||| |||||||| ||||||||||||||||||||||||||||
Db        113 AAGTAAGCTGCCCCCTTGTCCATACCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTG 54

Qy       1978 CCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                ||||| ||||| |||||| | |||||||||||||||||||||||||||||| |
Db          53 CCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCACC 1

RESULT 266
US-09-925-065A-296721/c
; Sequence 296721, Application US/09925065A
```

```
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 296721
;    LENGTH: 573
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-296721

  Query Match             6.3%;  Score 473.3;  DB 1;  Length 573;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  489;  Conservative   56;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:|||||||||| |:|||||||||||||:|||||||| ||:|:|:|||||:|||||||
Db        573 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 514

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||||:||||:||||||||||||||||:|||| ||||| ||
Db        513 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 454

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
              ||: :||| |||||||||||||||||||||||:|||||| |||||||||||||:|::
Db        453 AAACCTATCTAGCCTATACTGGCTTATCCTCATCCCAAAACACTAAAGCAACTAAGAGGG 394

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
             |||||||||:||| |||:||||||||||:||||||| |||||||:  ||:||||||||||
Db        393 TTCCTTGGCATAAAAGGCTTCTGCCGAATATGGATT-CCCAGGTACAGCGAAATAGCCAG 335

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
             |:|||||:|||||:||||||||||||||:|||||||||||||||||||:|||||||||:
Db        334 GCCATTATATACACTAATTAAGGAAACTCASAAAGCCAATACCCATTTAGTAAGATGGAT 275

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
             |:|||||:::|||||||||||||||||||  ||  | | |||  ||||||| :|||||||
Db        274 ACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTG 217

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
             |||||:|||||||:| || |||||||:||:|||::|||||  ||||||||:||:|||
Db        216 TTAAGCTTGCCAACAGTGCTAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGC 159

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
             |||:|||||||||||||:|| :|||||: ||||||||||::|||||:|| |||:|||
Db        158 TCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAA 99

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
             |||||:|||||||||| ||||||||||:|:||||||||||: ||||||||  ||||||||
Db          98 GGAAACTGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGTGGCAGCAGTAGC 39

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRA 3631
             |||::|||||||||||||||||||||| ||||||||:|
Db          38 AGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAA 1


RESULT 267
US-09-925-065A-296721/c
```

```
; Sequence 296721, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 296721
;    LENGTH: 573
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-296721

  Query Match             6.3%;  Score 473.3;  DB 1;  Length 573;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  489;  Conservative   56;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:|||||||||| |:|||||||||||||:|||||||| ||:|:|:|||||:|||||||
Db        573 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 514

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||||:||||:||||||||||||||||:|||| ||||| ||
Db        513 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 454

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
              ||: :||| |||||||||||||||||||||||:|||||| |||||||||||||:|::
Db        453 AAACCTATCTAGCCTATACTGGCTTATCCTCATCCCAAAACACTAAAGCAACTAAGAGGG 394

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
             |||||||||:||| |||:||||||||||:||||||| |||||||:  ||:||||||||||
Db        393 TTCCTTGGCATAAAAGGCTTCTGCCGAATATGGATT-CCCAGGTACAGCGAAATAGCCAG 335

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
             |:|||||:|||||:||||||||||||||:|||||||||||||||||||:|||||||||:
Db        334 GCCATTATATACACTAATTAAGGAAACTCASAAAGCCAATACCCATTTAGTAAGATGGAT 275

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
             |:|||||:::|||||||||||||||||||  ||  | | |||  ||||||| :|||||||
Db        274 ACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTG 217

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
             |||||:|||||||:| || |||||||:||:|||::|||||  ||||||||:||:|||
Db        216 TTAAGCTTGCCAACAGTGCTAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGC 159

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
             |||:|||||||||||||:|| :|||||: ||||||||||::|||||:|| |||:|||
Db        158 TCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAA 99

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
             |||||:|||||||||| ||||||||||:|:||||||||||: ||||||||  ||||||||
Db          98 GGAAACTGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGTGGCAGCAGTAGC 39

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRA 3631
             |||::|||||||||||||||||||||| ||||||||:|
Db          38 AGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAA 1


RESULT 268
```

```
US-09-925-065A-633131
; Sequence 633131, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633131
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-633131

  Query Match             6.3%;  Score 471;  DB 1;  Length 633;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  562;  Conservative    8;  Mismatches   54;  Indels    8;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
             |||||||| |||||||||||||||||||||:|||||||||||| ||:||||||| ||||
Db             9 AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCACTGGG 68

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                ||| | |     ||||| |||||||||| ||||||||||||||||||||  |||||
Db            69 ---ACACAGAATCAGAACAT-GGAGATTGGTTCTGCAGACATTTGCTAACTTGCATGCTA 124

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
                :|||||||||||||||:|||| ||: |||:|||||| ||||||||||||| |||||
Db           125 CAAGGACTAAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAATGATGTCCACTATAAC 183

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
                ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db           184 ACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGAA 242

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
                ||| | |||| |||||||||||||  |||||||||||||||||||||   |||||||
Db           243 AGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGAAAAGTTTA 302

Qy       1623 TCACTCAGTCAGCTGCAGACATTA-GAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
                | |||||||||||||||||||| ||||||||||||||||||  || ||||||||||||
Db           303 TTACTCAGTCAGCTGCAGACATTAGGAAAAAACTTCAAAAGTCCACCTTAGGCCCGGAGC 362

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
                |||||||||||||||||||||||||||||:||||||||||||||||:||||||||||
Db           363 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGMGATCAGGAGG 422

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
                | ||||| |||  |||||||| ||| |||||||||| |||||||||||||||| | |||
Db           423 AACAGGCAGAATGGGACAAACAGGA-TAAAAAAAAGACCACCGCTTTAGTCATGGCTCTC 481

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
                |||||||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||||
Db           482 AGGCAAGTGGACTTTGGAGGCTCTGGAAATGGGAAAAGCTAGGCAAATGGAATGCCTAAT 541

Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
                |||||||||||||||||| |||||   ||| |||||||||||||||||||||||| | |
Db           542 AGGGCTTGCTTCCAGTGCAGTCTATGGGGATACTTTAAAAAAGATTGTCCAAGTAGGAAT 601

Qy       1922 AAGCCGCCCCTTCGTCCATGCCCCTTATTTCA 1953
                ||||| | || | ||||| || |||||| |||.
Db           602 AAGCCACGCCCTTGTCCACGCTCCTTATGTCA 633


RESULT 269
US-09-925-065A-633131
; Sequence 633131, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633131
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-633131

  Query Match             6.3%;  Score 471;  DB 1;  Length 633;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  562;  Conservative    8;  Mismatches   54;  Indels    8;  Gaps    6;

Qy       1323 AGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGG 1382
             |||||||| |||||||||||||||||||||:|||||||||||| ||:||||||| ||||
Db             9 AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCACTGGG 68

Qy       1383 GACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTA 1442
                ||| | |     ||||| |||||||||| ||||||||||||||||||||  |||||
Db            69 ---ACACAGAATCAGAACAT-GGAGATTGGTTCTGCAGACATTTGCTAACTTGCATGCTA 124

Qy       1443 SAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAAC 1502
                :|||||||||||||||:|||| ||: |||:|||||| ||||||||||||| |||||
Db           125 CAAGGACTAAGGAAAACTAGGAAG-AAGCCTATGAATTATTCAATGATGTCCACTATAAC 183

Qy       1503 ACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGA 1562
                ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db           184 ACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGAA 242

Qy       1563 AGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTA 1622
                ||| | |||| |||||||||||||  |||||||||||||||||||||   |||||||
Db           243 AGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGAAAAGTTTA 302

Qy       1623 TCACTCAGTCAGCTGCAGACATTA-GAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGC 1681
                | |||||||||||||||||||| ||||||||||||||||||  || ||||||||||||
Db           303 TTACTCAGTCAGCTGCAGACATTAGGAAAAAACTTCAAAAGTCCACCTTAGGCCCGGAGC 362

Qy       1682 AAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGG 1741
                |||||||||||||||||||||||||||||:||||||||||||||||:||||||||||
Db           363 AAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGMGATCAGGAGG 422

Qy       1742 AGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTC 1801
                | ||||| |||  |||||||| ||| |||||||||| |||||||||||||||| | |||
Db           423 AACAGGCAGAATGGGACAAACAGGA-TAAAAAAAAGACCACCGCTTTAGTCATGGCTCTC 481

Qy       1802 AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAAT 1861
                |||||||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||||
Db           482 AGGCAAGTGGACTTTGGAGGCTCTGGAAATGGGAAAAGCTAGGCAAATGGAATGCCTAAT 541
```

```
Qy       1862 AGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGT 1921
Db        542 AGGGCTTGCTTCCAGTGCAGTCTATGGGGATACTTTAAAAAAGATTGTCCAAGTAGGAAT 601

Qy       1922 AAGCCGCCCCTTCGTCCATGCCCCTTATTTCA 1953
Db        602 AAGCCACGCCCTTGTCCACGCTCCTTATGTCA 633


RESULT 270
US-09-925-065A-926542/c
; Sequence 926542, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926542
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926542

  Query Match             6.3%;  Score 470.9;  DB 1;  Length 617;
  Best Local Similarity   83.6%;  Pred. No. 0;
  Matches  521;  Conservative   33;  Mismatches   60;  Indels    9;  Gaps    3;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        617 AATAGCTCTCACTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGC 558

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        557 CTCTAAGTATGCTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATT 498

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
Db        497 CCTAACTTCCAAGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCY 438

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
Db        437 TGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGA 383

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
Db        382 AAGGGAAGTAGAAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGA 324

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
Db        323 CCTTCCATTAGAAATGCTTATAGAAGGATGCCTAGTATAGGGTAATCCCCTCTGGGAAAC 264

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
Db        263 CAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCT 204

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
Db        203 CCCCTCCAGATGGCTAGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACA 144

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
Db        143 GAAATTACTTAAAACCCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGAT 84

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGC 4629
Db          83 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGG 24

Qy       4630 CTGTGAAKTGTGCCARARAAATA 4652
Db          23 CTGTGAAGTGTGCCAAAGAAATA 1


RESULT 271
US-09-925-065A-926542/c
; Sequence 926542, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926542
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926542

  Query Match             6.3%;  Score 470.9;  DB 1;  Length 617;
  Best Local Similarity   83.6%;  Pred. No. 0;
  Matches  521;  Conservative   33;  Mismatches   60;  Indels    9;  Gaps    3;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        617 AATAGCTCTCACTCAGGAACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATACAGC 558

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        557 CTCTAAGTATGCTTGCCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATT 498

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
Db        497 CCTAACTTCCAAGGGAACACCAGTCAAACATCAGGAAGCTATTAGGAGATTATAACTGCY 438

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
Db        437 TGTACAGAAACCTAAAGAGGTGACAGTCTTACACTGCCAGGGTCATCA-----GAAAGGA 383

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
Db        382 AAGGGAAGTAGAAAGGAACTGCCAAGTGGATATTGAAGTCAAAAGAGC-ACAAGGCAGGA 324

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
Db        323 CCTTCCATTAGAAATGCTTATAGAAGGATGCCTAGTATAGGGTAATCCCCTCTGGGAAAC 264

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
```

```
Db        263 CAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCT 204

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             ||||||  || || |||||||||| ||||| |||||||||  |||||| ||| ||||
Db        203 CCCCTCCAGATGGCTAGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACA 144

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             ||||||||||||||||||||| |||||||| |||||||||| ||||||||||||||:||
Db        143 GAAATTACTTAAAACCCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGAT 84

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGC 4629
             |||||||| |||||||||||||||||||||||||||||||||||| |:|||:||||||
Db         83 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGG 24

Qy       4630 CTGTGAAKTGTGCCARARAAATA 4652
             |||||||:|||||||:|:|||||
Db         23 CTGTGAAGTGTGCCAAAGAAATA 1


RESULT 272
US-10-301-480-280971
; Sequence 280971, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 280971
;   LENGTH: 623
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-280971

  Query Match             6.3%;  Score 470.8;  DB 1;  Length 623;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  522;  Conservative   30;  Mismatches   44;  Indels    8;  Gaps    4;

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
             || |||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| ||||||
Db          1 GATAAGGAAAAAGGGTAAATATATATGCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 60

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
             ||||||::||||||||:|:|||||||||||||||||| ||:|||:||||||||||||:||
Db         61 CCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACATCTGAGGGAACACCTATCAAAC 120

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
             | ||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:|||||
Db        121 ACCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 180

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: ||||||||| |||||||||:||||||||||:|||:|||: ||||||||
Db        181 TACACTGCCAGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCAA 240

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTC 4361
             ||||||||||:|||| |||  |||||| ||| ||| |||||| |||| |||| ||   |
Db        241 ATATTGAAGCCAAAAAAGCCACAAGGCGGGATCCTTCATTAGGAATGTTTATAGAAGGAC 300

Qy       4362 CCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAG 4421
             || |||||| |||||| ||||||   |||||||||||||||||||||||| ||||| ||
Db        301 CCCTAGTATGGGGTAA-CCCTTCTCAGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAG 359

Qy       4422 AATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAG 4478
             |||||||||||||| ||||    ||||| |||||||| ||| || |||||||  |||||
Db        360 AATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAA- 418

Qy       4479 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 4538
             |||||||||||||||||||| |||  |||||||||||||||||||||||||||||||||
Db        419 GGAAAATACTTTTGCCTGCAGCTAATCAATGGAAATTACTTAAAACCCTTCACCAAACCT 478

Qy       4539 TTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCC 4598
             |||||||||||||| |||||| |||||||:|||||||||||   ||||:|||||||||||
Db        479 TTCACTTAGGCATTGATAGCGCCCATCAGATGGCCAAATC---ATTTRCTGGACCAGGCC 535

Qy       4599 TTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCC 4658
             |||||||||||||||||||:|||:|||||||||||||| ||||||:|:|||||||||||
Db        536 TTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGAGTGCCAAAGAAATAATCCCC 595

Qy       4659 TGCC 4662
             ||||
Db        596 TGCC 599


RESULT 273
US-10-301-480-894380
; Sequence 894380, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 894380
;   LENGTH: 623
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-894380

  Query Match             6.3%;  Score 470.8;  DB 1;  Length 623;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  522;  Conservative   30;  Mismatches   44;  Indels    8;  Gaps    4;

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
             || |||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| ||||||
Db          1 GATAAGGAAAAAGGGTAAATATATATGCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 60

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
             ||||||::||||||||:|:|||||||||||||||||| ||:|||:||||||||||||:||
Db         61 CCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACATCTGAGGGAACACCTATCAAAC 120

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
             | ||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:|||||
Db        121 ACCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 180

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: ||||||||| |||||||||:||||||||||:|||:|||: ||||||||
Db        181 TACACTGCCAGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCAA 240

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTC 4361
             ||||||||||:|||| |||  |||||| ||| ||| |||||| |||| |||| ||   |
Db        241 ATATTGAAGCCAAAAAAGCCACAAGGCGGGATCCTTCATTAGGAATGTTTATAGAAGGAC 300

Qy       4362 CCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAG 4421
             || |||||| |||||| ||||||   |||||||||||||||||||||||| ||||| ||
Db        301 CCCTAGTATGGGGTAA-CCCTTCTCAGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAG 359

Qy       4422 AATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAG 4478
             |||||||||||||| ||||    ||||| |||||||| ||| || |||||||  |||||
Db        360 AATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAA- 418
```

```
Qy       4479 GGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCT 4538
             |||||||||||||||||||| ||| ||||||||||||||||||||||||||| |||||||
Db        419 GGAAAATACTTTTGCCTGCAGCTAATCAATGGAAATTACTTAAAACCCTTCACCAAACCT 478

Qy       4539 TTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCC 4598
             |||||||||||||| |||||| ||||||:||||||||||||   ||||:|||||||||||||
Db        479 TTCACTTAGGCATTGATAGCGCCCATCAGATGGCCAAATC---ATTTRCTGGACCAGGCC 535

Qy       4599 TTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCC 4658
             |||||||||||||||||:|||:|||||||||||||| |||||:|:||||||||||||
Db        536 TTTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGAGTGCCAAAGAAATAATCCCC 595

Qy       4659 TGCC 4662
             ||||
Db        596 TGCC 599
```

RESULT 274
US-09-925-065A-926543/c
; Sequence 926543, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926543
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926543

Query Match 6.3%; Score 470.2; DB 1; Length 560;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 487; Conservative 40; Mismatches 33; Indels 2; Gaps 2;

```
Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| ||| |||||:| |||||||||| |||||:||||||||||||||||||
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:|||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||| :|||    |||||||||| ||||:||: ||::|| ||||||||
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             ||||||||||||||||||||:|||||:||||| |||||:|||||||||||| |||||
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||| |::|:|:|||||:||||||||||||| ||||||||||||||:||||:||||||||
Db        201 CCAGAYGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||| |||||:||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||:|||||||||||||| |: ::|||||||||:|||:|||:||||||||||:|||||||
Db         81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
             |||||||:||  :||| |||||
Db         22 CCCAGGTATGATGAAACAGCCA 1
```

RESULT 275
US-09-925-065A-926543/c
; Sequence 926543, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926543
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926543

Query Match 6.3%; Score 470.2; DB 1; Length 560;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 487; Conservative 40; Mismatches 33; Indels 2; Gaps 2;

```
Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| ||| |||||:| |||||||||| |||||:||||||||||||||||||
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:|||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||| :|||    |||||||||| ||||:||: ||::|| ||||||||
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             ||||||||||||||||||||:|||||:||||| |||||:|||||||||||| |||||
```

```
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              ||| |::|:|:||||||:||||||||||||||| ||||||||||||||:||||:||||||||
Db        201 CCAGAYGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              ||| |||||:|||||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||||||||| |: ::|||||||||:|||:|||:||||||||||:|||||||
Db         81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
              |||||||:||  :||| |||||
Db         22 CCCAGGTATGATGAAACAGCCA 1


RESULT 276
US-09-925-065A-191356
; Sequence 191356, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 191356
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-191356

  Query Match             6.3%;  Score 469.8;  DB 1;  Length 600;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  521;  Conservative   30;  Mismatches   44;  Indels    8;  Gaps    4;

Qy       4063 AGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGC 4122
              | |||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| |||||||
Db          1 ATAAGGAAAAAGGGTAAATATATATGCAGACTCTAAGTATGCTTACCTAGTCCTCCATGC 60

Qy       4123 CCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACA 4182
              |||||::|||||||:|:|||||||||||||||||| ||:|||:||||||||||||:|||
Db         61 CCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACATCTGAGGGAACACCTATCAAACA 120

Qy       4183 TCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTT 4242
              ||||||||||||||||:|||||||:||||:||||||||||||:||||||||:|||||||
Db         121 CCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTT 180

Qy       4243 ACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKA 4302
              |||||||: ||||||||| |||||||||:||||||||||:|||:|||: |||||||| |
Db         181 ACACTGCCAGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCAAA 240

Qy       4303 TATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCC 4362
              |||||||||:|||| |||  |||||| ||| ||| |||||| |||| ||||  ||   ||
Db         241 TATTGAAGCCAAAAAAGCCACAAGGCGGGATCCTTCATTAGGAATGTTTATAGAAGGACC 300

Qy       4363 CTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGA 4422
              | |||||| |||||| |||||||   ||||||||||||||||||||||||||| ||||| |||
Db         301 CCTAGTATGGGGTAA-CCCTTCTCAGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGA 359

Qy       4423 ATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGG 4479
              |||||||||||| ||||    ||||| |||||||| ||| || ||||||| ||||| |
Db         360 ATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAA-G 418

Qy       4480 GAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTT 4539
              |||||||||||||||||| ||| ||||||||||||||||||||||||||||| ||||||||
Db         419 GAAAATACTTTTGCCTGCAGCTAATCAATGGAAATTACTTAAAACCCTTCACCAAACCTT 478

Qy       4540 TCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCT 4599
              |||||||||||| |||||| |||||||:||||||||||   ||||:|||||||||||||
Db         479 TCACTTAGGCATTGATAGCGCCCATCAGATGGCCAAATC---ATTTRCTGGACCAGGCCT 535

Qy       4600 TTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCT 4659
              |||||||||||||||||||:|||:|||||||||||||: ||||||:|:|||||||||||
Db         536 TTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGAGTGCCAAAGAAATAATCCCCT 595

Qy       4660 GCC 4662
              |||
Db         596 GCC 598


RESULT 277
US-09-925-065A-191356
; Sequence 191356, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 191356
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-191356

  Query Match             6.3%;  Score 469.8;  DB 1;  Length 600;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  521;  Conservative   30;  Mismatches   44;  Indels    8;  Gaps    4;

Qy       4063 AGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGC 4122
              | |||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| |||||||
Db          1 ATAAGGAAAAAGGGTAAATATATATGCAGACTCTAAGTATGCTTACCTAGTCCTCCATGC 60

Qy       4123 CCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACA 4182
              |||||::|||||||:|:|||||||||||||||||| ||:|||:||||||||||||:|||
Db         61 CCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACATCTGAGGGAACACCTATCAAACA 120

Qy       4183 TCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTT 4242
              ||||||||||||||||:|||||||:||||:||||||||||||:||||||||:|||||||
Db         121 CCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTT 180

Qy       4243 ACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKA 4302
              |||||||: ||||||||| |||||||||:||||||||||:|||:|||: |||||||| |
Db         181 ACACTGCCAGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGCAAA 240
```

```
Qy       4303 TATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCC 4362
             |||||||||:|||| |||  |||||| ||| ||| |||||| |||| ||||  ||   ||
Db        241 TATTGAAGCCAAAAAAGCCACAAGGCGGGATCCTTCATTAGGAATGTTTATAGAAGGACC 300

Qy       4363 CTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGA 4422
             | |||||| |||||| ||||||   ||||||||||||||||||||||||| ||||| |||
Db        301 CCTAGTATGGGGTAA-CCCTTCTCAGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGA 359

Qy       4423 ATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGG 4479
             ||||||||||||| ||||    |||| |||||||| ||| || |||||||  ||||| |
Db        360 ATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAA-G 418

Qy       4480 GAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTT 4539
             |||||||||||||||||| |||  |||||||||||||||||||||||||||||| ||||||||
Db        419 GAAAATACTTTTGCCTGCAGCTAATCAATGGAAATTACTTAAAACCCTTCACCAAACCTT 478

Qy       4540 TCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCT 4599
             |||||||||||| |||||| |||||||:|||||||||||   ||||:|||||||||||||
Db        479 TCACTTAGGCATTGATAGCGCCCATCAGATGGCCAAATC---ATTTRCTGGACCAGGCCT 535

Qy       4600 TTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCT 4659
             |||||||||||||||||:|||:|||||||||||||: |||||||:|:|||||||||||||
Db        536 TTTCAAAACTATCAAGCAGATAGTCAGGGCCTGTGAAGAGTGCCAAAGAAATAATCCCCT 595

Qy       4660 GCC 4662
             |||
Db        596 GCC 598


RESULT 278
US-09-925-065A-948890/c
; Sequence 948890, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948890
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948890

  Query Match             6.3%;  Score 469;  DB 1;  Length 560;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  488;  Conservative   38;  Mismatches   34;  Indels    2;  Gaps    2;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             |||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| |||  ||||||:| |||||||||| ||||||:||||||||||||||||||
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
             | ||||:||||||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||| :|||    |||||||| ||||:||: ||::|| ||||||||
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             |||||||||||||||||||||||:|||||:|||||  |||||:|||||||||||  |||||
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||| | :|:|:|||||:|||||||||||||| ||||||||||||||:||||:||||||||
Db        201 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||| |||||:|||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||||| |: ::|||||||||||||:|||:|||||||||||:|||||||
Db         81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCRTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
             |||||||:||  :||| |||||
Db         22 CCCAGGTATGATGAAACAGCCA 1


RESULT 279
US-09-925-065A-948890/c
; Sequence 948890, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948890
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948890

  Query Match             6.3%;  Score 469;  DB 1;  Length 560;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  488;  Conservative   38;  Mismatches   34;  Indels    2;  Gaps    2;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             |||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| |||  ||||||:| |||||||||| ||||||:||||||||||||||||||
```

```
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
              |||||||||||||||||| :|||    |||||||| ||||:||: ||::|| ||||||||
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
              |||||||||||||||||||||:|||||:|||||  |||||:|||||||||||||| |||||
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              ||| | :|:|:|||||:|||||||||||||| ||||||||||||||:||||:||||||||
Db        201 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              ||| |||||:|||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||||||||| |: ::|||||||||||||||:|||:||||||||||:|||||||
Db         81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCRTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
              |||||||:||  :||| |||||
Db         22 CCCAGGTATGATGAAACAGCCA 1


RESULT 280
US-09-925-065A-307095
; Sequence 307095, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 307095
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-307095

  Query Match             6.3%;  Score 468.9;  DB 1;  Length 601;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  522;  Conservative   24;  Mismatches   55;  Indels    9;  Gaps    3;

Qy       2501 CACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCA 2560
              ||| |||||||||||||||||||||||||||||| ||||   ||||| |||| ||| || |
Db          1 CACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCAA 60

Qy       2561 AAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAG 2620
              ||| |||||||||||| |||||||||||||| |||||| ||||||||  | |||||||
Db         61 AAAAATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCAA 120

Qy       2621 TAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAA 2680
              || | |        | |   |  | ||||||||||  |||||||||| ||||||||||
Db        121 TACTCC-------AACTTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAA 172

Qy       2681 GATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTA 2740
              |||||||||||||||||  ||||| |||||||  | ||||||||||||||||||  ||||||
Db        173 GATCTCAGGATTATCAAT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTA 231

Qy       2741 TACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGA 2800
              | | ||::|||||||||||||||||||||||||||||||||||:|||||||||||:||||
Db        232 CATTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGA 291

Qy       2801 TGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATAC 2860
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        292 TGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCC 351

Qy       2861 TTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCC 2920
              ||||||||||:  |||||||||||||||||:|||||||||||||||||||||||||:|
Db        352 TTCAAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCT 411

Qy       2921 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCT 2980
              |||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db        412 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCT 471

Qy       2981 TGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCA 3040
              ||||||||:|||: ||||||||||||||||:|   ||::|||||||||||||   |||
Db        472 TGTCCTTCAGTAYATGGATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGATGTCA 531

Qy       3041 AGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARA 3100
              |||||| |||||:|||||:| |||||||  :|||| ||||||||:||||| |||||:|:|
Db        532 AGCCACACAAGCGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTTTCAAACCAAA 591

Qy       3101 RGCTCARCTC 3110
              :|||||:|||
Db        592 GGCTCAGCTC 601


RESULT 281
US-09-925-065A-307095
; Sequence 307095, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 307095
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-307095

  Query Match             6.3%;  Score 468.9;  DB 1;  Length 601;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  522;  Conservative   24;  Mismatches   55;  Indels    9;  Gaps    3;

Qy       2501 CACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCA 2560
              ||| |||||||||||||||||||||||||||||| |||||   ||||| |||| ||| || |
Db          1 CACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACTTGAGGCCAAACAGGGACTCAA 60
```

```
Qy       2561 AAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAG 2620
Db         61 AAAAATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAACACCATGCAGTAGCCCCTGCAA 120

Qy       2621 TAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAA 2680
Db        121 TACTCC-------AACTTTAGCAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAA 172

Qy       2681 GATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTA 2740
Db        173 GATCTCAGGATTATCAAT-GAGGCTGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTA 231

Qy       2741 TACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGA 2800
Db        232 CATTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGA 291

Qy       2801 TGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATAC 2860
Db        292 TGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCC 351

Qy       2861 TTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCC 2920
Db        352 TTCAAACCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCT 411

Qy       2921 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCT 2980
Db        412 CCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCT 471

Qy       2981 TGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCA 3040
Db        472 TGTCCTTCAGTAYATGGATGATTTACTTTTAGTGACCCATTCAGAAACCTTGTGATGTCA 531

Qy       3041 AGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARA 3100
Db        532 AGCCACACAAGCGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTTTCAAACCAAA 591

Qy       3101 RGCTCARCTC 3110
Db         592 GGCTCAGCTC 601
```

RESULT 282
US-09-925-065A-926544/c
; Sequence 926544, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926544
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926544

Query Match 6.3%; Score 468.6; DB 1; Length 560;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 486; Conservative 40; Mismatches 34; Indels 2; Gaps 2;

```
Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCYACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db        201 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db         81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
Db         22 CCCAGGTATGATGAAACAGCCA 1
```

RESULT 283
US-09-925-065A-926544/c
; Sequence 926544, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926544
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926544

Query Match 6.3%; Score 468.6; DB 1; Length 560;
Best Local Similarity 86.5%; Pred. No. 0;

```
  Matches  486;  Conservative   40;  Mismatches   34;  Indels    2;  Gaps    2;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||||:| ||||
Db        560 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 501

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| ||||||||||||||||||||||||||||||||||   |||||||
Db        500 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 441

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| ||| |||||||:| |||||||||| |||||||:||||||||||||||||||
Db        440 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 381

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:|||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        380 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 321

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||||| :|||    ||||||||| ||||:||: ||::|| |||||||||
Db        320 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 261

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             |||||||||||||||||||||:|||||:||||| |||||:||||||||:||| |||||
Db        260 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCYACA-GGTTT 202

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||| | :|:|:|||||:|||||||||||||| ||||||||||||||:||||:|||||||||
Db        201 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 142

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||| |||||:|||||||||||||:::|||||||||||||||| |||||| |||:||
Db        141 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 82

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||| |: ::|||||||||:|||:|||:||||||||||:|||||||
Db          81 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATT- 23

Qy       3271 CCCAGGTWTGGCRAAATAGCCA 3292
             |||||||:||  :||| |||||
Db          22 CCCAGGTATGATGAAACAGCCA 1


RESULT 284
US-09-925-065A-205418
; Sequence 205418, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205418
;   LENGTH: 603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-205418

  Query Match             6.3%;  Score 468.4;  DB 1;  Length 603;

  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  521;  Conservative   30;  Mismatches   48;  Indels    6;  Gaps    4;

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
             |||:|||||||||| ||||||||||||||||||||||||:||||:||| |||||| ||||
Db          1 ATACTTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCC 60

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRA 4098
              :|||:| |||||||||||||||||||| |:||||||||:|||||||||||||||||:|
Db         61 CTCACTCGGGCACTAGAATTAGGAGAAGGAAAAAAGGGTAAATATATATACAGACTCTAA 120

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
              :||||||:|||||||| |||||||||||::|||||||:|:|||||||||||||||||
Db         121 GTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAAT 180

Qy       4159 TTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACA 4218
              ||| |||:|||||||||||:|||||||||||||||||||:|||||||:||||:|||||
Db         181 TTC-GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACA 239

Qy       4219 GAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGA 4278
              |:||||| :|||||||:|||||||||||||:    |||||| |||||||||:|||||||
Db         240 GRAACCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGA 299

Qy       4279 AATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCC 4338
              ||||:|||:|||: |||||||||:||||||||| :|||||||||||||||||||||||||
Db         300 AATAGAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCC 359

Qy       4339 ATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCC 4398
              |||||||||||||||  ||   ||| |||||| |||||||||  ||||||||||||| |
Db         360 ATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTC 419

Qy       4399 CCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTC 4455
              |||||||||||||||| ||| ||||| ||||||||||||||    | ||  ||||||||
Db         420 CCAGTACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTC 479

Qy       4456 GGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATT 4515
                || || |||||||   |||||| |||||| |||||||||||| ||| ||||| ||||||
Db         480 CAGATGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATT 538

Qy       4516 ACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAA 4575
              |||||||||||||||| | |||||| |||||||||| ||||| |||||||:||||||||
Db         539 ACTTAAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAA 598

Qy       4576 ATCAT 4580
              || ||
Db         599 ATTAT 603


RESULT 285
US-09-925-065A-205418
; Sequence 205418, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205418
;   LENGTH: 603
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-205418

  Query Match             6.3%;  Score 468.4;  DB 1;  Length 603;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  521;  Conservative   30;  Mismatches   48;  Indels    6;  Gaps    4;

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||| |||||||||||||||||||||||:||||:||| |||||| ||||
Db             1 ATACTTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCC 60

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRA 4098
              :|||:| |||||||||||||||||||| |:||||||||:||||||||||||||||||:|
Db            61 CTCACTCGGGCACTAGAATTAGGAGAAGGAAAAAAGGGTAAATATATATACAGACTCTAA 120

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
              :||||||:|||||||| |||||||||||||::|||||||:|:|||||||||||||||||
Db          121 GTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAAT 180

Qy       4159 TTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACA 4218
              ||| |||:||||||||||||:||||||||||||||||||||:|||||||:||||:|||||
Db          181 TTC-GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACA 239

Qy       4219 GAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGA 4278
              |:|||||| :|||||||:|||||||||||||:    |||||| |||||||||:|||||||
Db          240 GRAACCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGA 299

Qy       4279 AATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCC 4338
              ||||:|||:|||: |||||||:|||||||| :|||||||||||||||||||||||||||
Db          300 AATAGAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCC 359

Qy       4339 ATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCC 4398
              |||||||||||||| ||   ||| |||||| |||||||| ||||||||||||||| |
Db          360 ATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTC 419

Qy       4399 CCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTC 4455
              |||||||||||||| ||| ||||| |||||||||||||||    | ||  ||||||||
Db          420 CCAGTACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTC 479

Qy       4456 GGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATT 4515
                || || |||||||   |||||| |||||| |||||||||||| ||| ||||| ||||||
Db          480 CAGATGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATT 538

Qy       4516 ACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAA 4575
              |||||||||||||| | |||||| |||||||||||| ||||| ||||||||:||||||||
Db          539 ACTTAAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAA 598

Qy       4576 ATCAT 4580
              || ||
Db          599 ATTAT 603


RESULT 286
US-10-717-580-49
; Sequence 49, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (84)..(84)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (193)..(193)
;   OTHER INFORMATION: n = any nucleotide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (241)..(241)
;   OTHER INFORMATION: n = any nucleotide
US-10-717-580-49

  Query Match             6.3%;  Score 468.1;  DB 1;  Length 486;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  475;  Conservative   11;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        241 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGGCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||||:||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGG 479

Qy        595 ACATCCM 601
              |||||||
Db        480 ACATCCA 486


RESULT 287
US-09-925-065A-419320
; Sequence 419320, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 419320
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-419320

  Query Match             6.2%;  Score 467.7;  DB 1;  Length 588;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  504;  Conservative   27;  Mismatches   52;  Indels    5;  Gaps    2;

Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGA 4092
             |||||:|||:| |||||||| |||||||| |||:||||| ||:||  ||||||||||||
Db          1 TAGCCCTCACTCGGGCACTACAATTAGGAAAAGGAAAAAAGGTAAATATATATATACAGA 60

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| |||||||||||::|||||||:|:||||||||||||
Db         61 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCGGCAATATGGAGAGAAAGGGAATT 120

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||:|||:||||||||||||:|||||||||||||||||||:|||| ||:||||
Db        121 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTACTATTGGC 180

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
             :|||||||||||||:||||||||:|||||||||| ||: |||||| || |||| ||||:|
Db        181 TGTACAGAAACCTAAAGAGGTGGCAGTCTTACACGGCCTGGGTCACCAGAAAGAAAAGGA 240

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             |||||||||| || :|||:  |||||| :||||||||||:|||||||||||||||||||||
Db        241 AAGGGAAATAAAAAGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGGCAGGA 300

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
             |||||||||||||||||||||  ||    |   || ||| |||||||| | ||| ||||||
Db        301 CCCTCCATTAGAAATGCTTATAGAAGGACATCTACTATGGGGTAATCTCCTCCAGGAAAC 360

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
             |||||||||||||||| || ||||| |||||| |||| |||| || |    ||||   |
Db        361 CAAGCCCCAGTACTCAGAAGAAGAAATAGAATGAGGAAACTCATGAAGACATAGTTCCTT 420

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             ||:||| ||| || |||||||  ||||||| |||||||||||||||||| ||| ||||||
Db        421 CCYCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAGCCAATG 480

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             |||||||||||||| |||||| |||||||||||||||||| ||||||||||||||:||
Db        481 GAAATTACTTAAAATCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGAT 540

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 4617
             |||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 588


RESULT 288
US-09-925-065A-419320
; Sequence 419320, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 419320
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-419320

  Query Match             6.2%;  Score 467.7;  DB 1;  Length 588;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  504;  Conservative   27;  Mismatches   52;  Indels    5;  Gaps    2;

Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGA 4092
             |||||:|||:| |||||||| |||||||| |||:||||| ||:||  ||||||||||||
Db          1 TAGCCCTCACTCGGGCACTACAATTAGGAAAAGGAAAAAAGGTAAATATATATATACAGA 60

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| |||||||||||::|||||||:|:||||||||||||
Db         61 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCGGCAATATGGAGAGAAAGGGAATT 120

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||:|||:||||||||||||:|||||||||||||||||||:|||| ||:||||
Db        121 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTACTATTGGC 180

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
             :|||||||||||||:||||||||:|||||||||| ||: |||||| || |||| ||||:|
Db        181 TGTACAGAAACCTAAAGAGGTGGCAGTCTTACACGGCCTGGGTCACCAGAAAGAAAAGGA 240

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             |||||||||| || :|||:  |||||| :||||||||||:|||||||||||||||||||||
Db        241 AAGGGAAATAAAAAGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGGCAGGA 300

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
             |||||||||||||||||||||  ||    |   || ||| |||||||| | ||| ||||||
Db        301 CCCTCCATTAGAAATGCTTATAGAAGGACATCTACTATGGGGTAATCTCCTCCAGGAAAC 360

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
             |||||||||||||||| || ||||| |||||| |||| |||| || |    ||||   |
Db        361 CAAGCCCCAGTACTCAGAAGAAGAAATAGAATGAGGAAACTCATGAAGACATAGTTCCTT 420

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             ||:||| ||| || |||||||  ||||||| |||||||||||||||||| ||| ||||||
Db        421 CCYCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAGCCAATG 480

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             |||||||||||||| |||||| |||||||||||||||||| ||||||||||||||:||
Db        481 GAAATTACTTAAAATCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGAT 540

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 4617
             |||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 588


RESULT 289
US-09-925-065A-925327/c
; Sequence 925327, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
```

```
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925327
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925327

  Query Match             6.2%;  Score 467.3;  DB 1;  Length 552;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  473;  Conservative   38;  Mismatches   30;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| ||| ||||||:| |||||||||| ||||||:|||||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:|||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||| :|||    ||||||||| ||||:||: ||::|| ||||||||
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             |||||||||||||||||||||:|||||:|||||  |||||:|||||||||||||| |||||
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||| |::|:|:|||||:|||||||||||||| ||||||||||||||:||||:||||||||
Db        193 CCAGAYGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||| |||||:|||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||||  |: ::|||||||||:|||:|||:||||||||||:||||||||
Db          73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
              ||
Db          13 CC 12


RESULT 290
US-09-925-065A-925327/c
; Sequence 925327, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
```

```
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925327
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925327

  Query Match             6.2%;  Score 467.3;  DB 1;  Length 552;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  473;  Conservative   38;  Mismatches   30;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||||||||||||||| ||::|||||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             || ||||| ||| ||||||:| |||||||||| ||||||:|||||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              | ||||:|||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||| :|||    ||||||||| ||||:||: ||::|| ||||||||
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             |||||||||||||||||||||:|||||:|||||  |||||:|||||||||||||| |||||
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             ||| |::|:|:|||||:|||||||||||||| ||||||||||||||:||||:||||||||
Db        193 CCAGAYGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||| |||||:|||||||||||||||:::|||||||||||||||||| |||||| |||:||
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||||  |: ::|||||||||:|||:|||:||||||||||:||||||||
Db          73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
              ||
Db          13 CC 12


RESULT 291
US-09-925-065A-205417
; Sequence 205417, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205417
;   LENGTH: 603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-205417

  Query Match             6.2%;  Score 467.2;  DB 1;  Length 603;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  522;  Conservative   28;  Mismatches   49;  Indels    6;  Gaps    4;

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
             |||:|||||||||| |||||||||||||||||||||||:||||:||| |||||| ||||
Db             1 ATACTTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCC 60

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRA 4098
             :|||:| |||||||||||||||||||| |||||||||:|||||||||||||||||||:|
Db            61 CTCACTCGGGCACTAGAATTAGGAGAAGGRAAAAAGGGTAAATATATATACAGACTCTAA 120

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
             :|||||:|||||||| ||||||||||||::|||||||:|:|||||||||||||||||||
Db           121 GTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAAT 180

Qy       4159 TTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACA 4218
             ||| |||:|||||||||||:|||||||||||||||||||:||||||:||||:|||||
Db           181 TTC-GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACA 239

Qy       4219 GAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGA 4278
             | ||||| :||||||||:|||||||||||||:    |||||| |||||||||:|||||||
Db           240 GGAACCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGA 299

Qy       4279 AATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCC 4338
             ||||:|||:|||: |||||||||:|||||||| :|||||||||||||||||||||||||||
Db           300 AATAGAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCC 359

Qy       4339 ATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCC 4398
             |||||||||||||| ||   ||| |||||| |||||||| ||||||||||||||||| |
Db           360 ATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTC 419

Qy       4399 CCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTC 4455
             |||||||||||||||| ||| ||||| ||||||||||||||    | ||  ||||||||
Db           420 CCAGTACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTC 479

Qy       4456 GGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATT 4515
               || || |||||||   |||||| |||||| |||||||||||| ||| ||||| ||||||
Db           480 CAGATGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATT 538

Qy       4516 ACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAA 4575
             |||||||||||||| | |||||| |||||||||| ||||| ||||||||:||||||||
Db           539 ACTTAAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAA 598

Qy       4576 ATCAT 4580
             || ||
Db           599 ATTAT 603

RESULT 292
```

```
US-09-925-065A-205417
; Sequence 205417, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205417
;   LENGTH: 603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-205417

  Query Match             6.2%;  Score 467.2;  DB 1;  Length 603;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  522;  Conservative   28;  Mismatches   49;  Indels    6;  Gaps    4;

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
             |||:|||||||||| |||||||||||||||||||||||:||||:||| |||||| ||||
Db             1 ATACTTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCC 60

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRA 4098
             :|||:| |||||||||||||||||||| |||||||||:|||||||||||||||||||:|
Db            61 CTCACTCGGGCACTAGAATTAGGAGAAGGRAAAAAGGGTAAATATATATACAGACTCTAA 120

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
             :|||||:|||||||| ||||||||||||::|||||||:|:|||||||||||||||||||
Db           121 GTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAAT 180

Qy       4159 TTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACA 4218
             ||| |||:|||||||||||:|||||||||||||||||||:||||||:||||:|||||
Db           181 TTC-GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACA 239

Qy       4219 GAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGA 4278
             | ||||| :||||||||:|||||||||||||:    |||||| |||||||||:|||||||
Db           240 GGAACCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGA 299

Qy       4279 AATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCC 4338
             ||||:|||:|||: |||||||||:|||||||| :|||||||||||||||||||||||||||
Db           300 AATAGAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCC 359

Qy       4339 ATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCC 4398
             |||||||||||||| ||   ||| |||||| |||||||| ||||||||||||||||| |
Db           360 ATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTC 419

Qy       4399 CCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTC 4455
             |||||||||||||||| ||| ||||| ||||||||||||||    | ||  ||||||||
Db           420 CCAGTACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTC 479

Qy       4456 GGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATT 4515
               || || |||||||   |||||| |||||| |||||||||||| ||| ||||| ||||||
Db           480 CAGATGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATT 538

Qy       4516 ACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAA 4575
             |||||||||||||| | |||||| |||||||||| ||||| ||||||||:||||||||
Db           539 ACTTAAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAA 598

Qy       4576 ATCAT 4580
```

```
                 || ||
Db        599 ATTAT 603


RESULT 293
US-10-027-632-213458/c
; Sequence 213458, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 213458
;   LENGTH: 613
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-213458

  Query Match             6.2%;  Score 466.9;  DB 1;  Length 613;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  491;  Conservative   29;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4090 AGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGA 4149
             |||||||:|:||||||:|||||||| |||||||||  ::|||||||:|:|||||||||
Db        613 AGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGGA 554

Qy       4150 ATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYT 4209
             ||||||||||||:|||:||||||||||:||| ||||||||||||||:||||||||:|
Db        553 ATTCCTAACTTCTGAGGGAACACCTATCAAACAACAGGAAGCCATTAGGAGATTATTATT 494

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
             |||:|||||||||||||:||||||||:|||||||||||||||:|||||||||| ||||||||
Db        493 GGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 434

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
             |:||||||||||| |||:|||:  |||||| :|||||  ||||:|||| |||  |||||||
Db        433 GGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAGGCA 374

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
             |||| |||||||||||||||||  ||   ||| ||  || |||||||||| |||||||
Db        373 GGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTATCATGGGGTAATCCCCTCCGGGA 314

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTT 4446
             |||||||:|||||||||||||||| ||||| || |||||||| |||||||  |    |||||
Db        313 AACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGATGACATAGTTT 254

Qy       4447 TCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
             ||||||||  || || |||||||  |||||| |||||||| ||||||||  ||| |||
Db        253 CCTCCCCTCATGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAACCA 194

Qy       4507 ATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCA 4566
             || ||||||||||||||||||| ||||| |||||||| |||| ||||||||||||||||
Db        193 GTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTATGCATTGATAGCACCCATCA 134
```

```
Qy       4567 RATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAG 4626
             :|||||||||||||||||||| ||||||||||||||||||||||||||||:|||:||||
Db        133 GATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAG 74

Qy       4627 GGCCTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |:||| ||||:|||||||:|:||| || |||||||
Db          73 GRCCTATGAAGTGTGCCAAAGAAAGAACCCCCTGC 39


RESULT 294
US-10-027-632-213458/c
; Sequence 213458, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 213458
;   LENGTH: 613
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-213458

  Query Match             6.2%;  Score 466.9;  DB 1;  Length 613;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  491;  Conservative   29;  Mismatches   52;  Indels    3;  Gaps    1;

Qy       4090 AGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGA 4149
             |||||||:|:||||||:|||||||| |||||||||  ::|||||||:|:|||||||||
Db        613 AGACTCTAAGTATGCTTACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGGA 554

Qy       4150 ATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYT 4209
             ||||||||||||:|||:||||||||||:||| ||||||||||||||:||||||||:|
Db        553 ATTCCTAACTTCTGAGGGAACACCTATCAAACAACAGGAAGCCATTAGGAGATTATTATT 494

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
             |||:|||||||||||||:||||||||:|||||||||||||||:|||||||||| ||||||||
Db        493 GGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 434

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
             |:||||||||||| |||:|||:  |||||| :|||||  ||||:|||| |||  |||||||
Db        433 GGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAGGCA 374

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
             |||| |||||||||||||||||  ||   ||| ||  || |||||||||| |||||||
Db        373 GGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTATCATGGGGTAATCCCCTCCGGGA 314

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTT 4446
             |||||||:|||||||||||||||| ||||| || |||||||| |||||||  |    |||||
Db        313 AACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGATGACATAGTTT 254

Qy       4447 TCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
             ||||||||  || || |||||||  |||||| |||||||| ||||||||  ||| |||
Db        253 CCTCCCCTCATGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAACCA 194
```

```
Qy       4507 ATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCA 4566
              || |||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        193 GTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTATGCATTGATAGCACCCATCA 134

Qy       4567 RATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAG 4626
              :|||||||||||||||||||| |||||||||||||||||||||||||||:|||:||||
Db        133 GATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGTCAG 74

Qy       4627 GGCCTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
              |:||| ||||:|||||||:|:||| || |||||||
Db          73 GRCCTATGAAGTGTGCCAAAGAAAGAACCCCCTGC 39


RESULT 295
US-09-925-065A-775883
; Sequence 775883, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 775883
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-775883

  Query Match             6.2%;  Score 466.9;  DB 1;  Length 614;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  525;  Conservative   11;  Mismatches   57;  Indels   21;  Gaps    2;

Qy       6893 GGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTA 6952
              ||| |||||||||||:|||| || |||||||||| || ||||||||||||||||||| ||
Db          1 GGGACCTCCTCAGCCAATGGGTGTCCTGGATTCTTCCTTTCTTAGGACCTCTAGCAGTTA 60

Qy       6953 TAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTT 7012
              |||||||| |||| |||||||||||||| ||||||:|||| |||||||| ||||||||||
Db         61 TAATATTGTTACTTCTCTTTGGACCCTGCATCTTTAACCTGCTTGTTAAGTTTGTCTCTT 120

Qy       7013 CCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCA 7060
              ||||||| ||||||||:|| |||          |||||||||||| ||| ||||||||
Db        121 CCAGAATTGAAGCTGTAAAGCTACAAATGATTCTTCAAATGGAGCCCCAGAGGCAGTCCA 180

Qy       7061 AGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATG 7120
              |||||| |||||| || ||||| | ||||||||||: |||||| | ||  ||||| |||
Db        181 TGACTAAAATCTACTGCGGACCCTTAGACCGGCCTGCCAGCCCATGCTCCAATGTTGATG 240

Qy       7121 ACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATT 7180
              ||||| |||| |||||||| |||||||||| |||||||||  |||||  |||||||||
Db        241 ACATCCAAGGTACCCCTCCGGAGGAAATCTCAACTGCACAACTCCTACTGTGCCCCAATT 300

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
              ||||||||||||||||||||: ||| |  ||||||||||||||||||||||||||||||
Db        301 CAGCAGGAAGCAGTTAGAGCGGTCATCGACAGACCTCCCCAACAGCACTTAGGTTTTCCT 360

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
              | ||||| |||| |||||||||| ||||||| |||||||||||||| :|:||||||||||

Db        361 GATGAGAGGGGGAACTGAGAGACGGGACTAGTTGGATTTCCTAGGCCRACTAAGAATCCA 420

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
              |||||||||:|||||||||||| ||||||||||||||||||||||||||||||||||:||
Db        421 TAAGCCTAGCTGGGAAGGTGACAGCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 480

Qy       7361 ACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGG 7411
              ||||| ||||||||          |||||||||||||| |||||| ||||| || |||||
Db        481 ACACCAGACCAATCAGGTAGTAAAGAGAGCTCACTAAACTGCTAACTAGGCTAAAACAGG 540

Qy       7412 AGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGG 7471
              |||||||||||||||||||||:|||  |:|||||||||||| ||||||||||||||||||
Db        541 AGGTAAAGAAATAGCCAATCATCTATCTCCTGAGAGCACAGCGGGAGGGACAATGATCGG 600

Qy       7472 GATATAAACCCAAG 7485
              ||||||||||||| |
Db        601 GATATAAACCCAGG 614


RESULT 296
US-09-925-065A-775883
; Sequence 775883, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 775883
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-775883

  Query Match             6.2%;  Score 466.9;  DB 1;  Length 614;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  525;  Conservative   11;  Mismatches   57;  Indels   21;  Gaps    2;

Qy       6893 GGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTA 6952
              ||| |||||||||||:|||| || |||||||||| || ||||||||||||||||||| ||
Db          1 GGGACCTCCTCAGCCAATGGGTGTCCTGGATTCTTCCTTTCTTAGGACCTCTAGCAGTTA 60

Qy       6953 TAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTT 7012
              |||||||| |||| |||||||||||||| ||||||:|||| |||||||| ||||||||||
Db         61 TAATATTGTTACTTCTCTTTGGACCCTGCATCTTTAACCTGCTTGTTAAGTTTGTCTCTT 120

Qy       7013 CCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCA 7060
              ||||||| |||||||||:|| |||          |||||||||||| ||| ||||||||
Db        121 CCAGAATTGAAGCTGTAAAGCTACAAATGATTCTTCAAATGGAGCCCCAGAGGCAGTCCA 180

Qy       7061 AGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATG 7120
              |||||| |||||| || ||||| | ||||||||||: |||||| | ||  ||||| |||
Db        181 TGACTAAAATCTACTGCGGACCCTTAGACCGGCCTGCCAGCCCATGCTCCAATGTTGATG 240

Qy       7121 ACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATT 7180
              ||||| |||| |||||||| |||||||||| |||||||||  |||||  |||||||||
Db        241 ACATCCAAGGTACCCCTCCGGAGGAAATCTCAACTGCACAACTCCTACTGTGCCCCAATT 300

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
```

```
                |||||||||||||||||||||||: ||| |  ||||||||||||||||||||||||||||||
Db        301 CAGCAGGAAGCAGTTAGAGCGGTCATCGACAGACCTCCCCAACAGCACTTAGGTTTTCCT 360

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
                | ||||| |||| |||||||||| |||||||| |||||||||||||| :|:|||||||||
Db        361 GATGAGAGGGGGAACTGAGAGACGGGACTAGTTGGATTTCCTAGGCCRACTAAGAATCCA 420

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
                |||||||||:|||||||||||||  ||||||||||||||||||||||||||||||:||
Db        421 TAAGCCTAGCTGGGAAGGTGACAGCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTC 480

Qy       7361 ACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGG 7411
                |||||  ||||||||         |||||||||||||||| ||||||  |||||  ||  |||||
Db        481 ACACCAGACCAATCAGGTAGTAAAGAGAGCTCACTAAACTGCTAACTAGGCTAAAACAGG 540

Qy       7412 AGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGG 7471
                |||||||||||||||||||||:|||  |:||||||||||||| ||||||||||||||||||
Db        541 AGGTAAAGAAATAGCCAATCATCTATCTCCTGAGAGCACAGCGGGAGGGACAATGATCGG 600

Qy       7472 GATATAAACCCAAG 7485
                |||||||||||| |
Db        601 GATATAAACCCAGG 614


RESULT 297
US-09-925-065A-948281/c
; Sequence 948281, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948281
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948281

  Query Match             6.2%;  Score 466.1;  DB 1;  Length 552;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  474;  Conservative   36;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
                |||||||||||||| ||::|||||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
                |||||:|||||||||| |||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
                || ||||| |||  ||||||:| |||||||||||| ||||||:||||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
                | ||||:||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313
```

```
Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
                |||||||||||||||||| :|||    |||||||||| ||||:||: ||::|| ||||||||
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
                |||||||||||||||||||||:|||||:||||| |||||:|||||||||||| |||||
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
                ||| | :|:|:|||||:|||||||||||||| |||||||||||||:||||:||||||||
Db        193 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
                ||| |||||:|||||||||||||:::|||||||||||||||| |||||| |||:||
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
                |||||||||||||||| |: ::|||||||||||||:|||:||||||||||:||||||||
Db         73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCRTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
                ||
Db         13 CC 12


RESULT 298
US-09-925-065A-948281/c
; Sequence 948281, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948281
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948281

  Query Match             6.2%;  Score 466.1;  DB 1;  Length 552;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  474;  Conservative   36;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
                |||||||||||||| ||::|||||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
                |||||:|||||||||| |||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
                || ||||| |||  ||||||:| |||||||||||| ||||||:||||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
                | ||||:||||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313
```

```
Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
             |||||||||||||||||  :|||    |||||||||  ||||:||: ||::|| ||||||||
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
             ||||||||||||||||||||:|||||:|||||  |||||:|||||||||||||  |||||
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             |||  | :|:|:|||||:|||||||||||||||  ||||||||||||||:||||:||||||||
Db        193 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             |||  |||||:|||||||||||||||:::|||||||||||||||||  ||||||  |||:||
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||||  |: ::|||||||||||||:|||:|||||||||:||||||||
Db          73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCRTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
             ||
Db          13 CC 12
```

RESULT 299
US-09-925-065A-419319
; Sequence 419319, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 419319
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-419319

Query Match 6.2%; Score 466.1; DB 1; Length 588;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 503; Conservative 27; Mismatches 53; Indels 5; Gaps 2;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGA 4092
             |||||:|||:| |||||||| |||||||| |||:||||| ||:||  ||||||||||||
Db          1 TAGCCCTCACTCGGGCACTACAATTAGGAAAAGGAAAAAAGGTAAATATATATATACAGA 60

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| |||||||||:||::|||||||:|:||||||||||||
Db         61 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCWTGCGGCAATATGGAGAGAAAGGGAATT 120

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||:|||:||||||||||||:|||||||||||||||||||:|||| ||:||||
Db        121 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTACTATTGGC 180

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
             :|||||||||||||:||||||||:|||||||||| ||: |||||| || |||| ||||:|
Db        181 TGTACAGAAACCTAAAGAGGTGGCAGTCTTACACGGCCTGGGTCACCAGAAAGAAAAGGA 240

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
             |||||||||| || :|||:  |||||| :||||||||||:|||||||||||||||||||||
Db        241 AAGGGAAATAAAAAGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGGCAGGA 300

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
             ||||||||||||||||||||  ||   |  || ||| |||||||| | ||| ||||||
Db        301 CCCTCCATTAGAAATGCTTATAGAAGGACATCTACTATGGGGTAATCTCCTCCAGGAAAC 360

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
             |||||||||||||||||  || |||||  ||||||  |||| |||| || |    ||||  |
Db        361 CAAGCCCCAGTACTCAGAAGAAGAAATAGAATGAGGAAACTCATGAAGACATAGTTCCTT 420

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
             || ||| ||| || |||||||  |||||||  ||||||||||||||||||| ||| ||||||
Db        421 CCTCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAGCCAATG 480

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
             ||||||||||||||  ||||||  ||||||||||||||||||||  |||||||||||||:||
Db        481 GAAATTACTTAAAATCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGAT 540

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 4617
             |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 588
```

RESULT 300
US-09-925-065A-419319
; Sequence 419319, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 419319
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-419319

Query Match 6.2%; Score 466.1; DB 1; Length 588;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 503; Conservative 27; Mismatches 53; Indels 5; Gaps 2;

```
Qy       4035 TAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGA 4092
             |||||:|||:| |||||||| |||||||| |||:||||| ||:||  ||||||||||||
Db          1 TAGCCCTCACTCGGGCACTACAATTAGGAAAAGGAAAAAAGGTAAATATATATATACAGA 60

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:|||||||:|||||||| |||||||||:||::|||||||:|:||||||||||||
Db         61 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCWTGCGGCAATATGGAGAGAAAGGGAATT 120

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGC 4212
             |||||||||:|||:||||||||||||:|||||||||||||||||||:|||| ||:||||
Db        121 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTACTATTGGC 180

Qy       4213 WGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRA 4272
```

```
                :|||||||||||||:||||||||:|||||||||| ||: |||||| || |||| ||||:|
Db        181 TGTACAGAAACCTAAAGAGGTGGCAGTCTTACACGGCCTGGGTCACCAGAAAGAAAAGGA 240

Qy       4273 AAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGA 4332
                |||||||||| || :|||:  |||||| :||||||||||:||||||||||||||||||
Db        241 AAGGGAAATAAAAAGGAACCACCAAGCGGATATTGAAGCCAAAAGAGCTGCAAGGCAGGA 300

Qy       4333 CCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAAC 4392
                |||||||||||||||||||  ||    |   || ||| |||||||| | ||| ||||||
Db        301 CCCTCCATTAGAAATGCTTATAGAAGGACATCTACTATGGGGTAATCTCCTCCAGGAAAC 360

Qy       4393 CAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCT 4449
                |||||||||||||||| || ||||| |||||| |||| |||| || |    ||||   |
Db        361 CAAGCCCCAGTACTCAGAAGAAGAAATAGAATGAGGAAACTCATGAAGACATAGTTCCTT 420

Qy       4450 CCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATG 4509
                || ||| ||| || |||||||  ||||||| |||||||||||||||||| ||| ||||||
Db        421 CCTCTCAGGATGGCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAGCCAATG 480

Qy       4510 GAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARAT 4569
                |||||||||||||| |||||| ||||||||||||||||||| ||||||||||||||:||
Db        481 GAAATTACTTAAAATCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATCAGAT 540

Qy       4570 GGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 4617
                |||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 GGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCA 588


RESULT 301
US-09-925-065A-925328/c
; Sequence 925328, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925328
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925328

  Query Match             6.2%;  Score 465.7;  DB 1;  Length 552;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  472;  Conservative   38;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
                |||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
                |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
                || ||||| |||  ||||||:| |||||||||| ||||||:|||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
                | ||||:|||||||||||||||||||||||||||||||||||||:||: ::|||||||
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
                |||||||||||||||||  :|||    |||||||| ||||:||: ||::|| ||||||||
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
                |||||||||||||||||||||:|||||:||||| |||||:||||||||:||| |||||
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCYACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
                ||| | :|:|:|||||:|||||||||||||| ||||||||||||||:||||:|||||||
Db        193 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
                ||| |||||:|||||||||||||||:::|||||||||||||||| |||||| |||:||
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
                |||||||||||||||||  |: ::|||||||||:|||:|||:|||||||||||:||||||||
Db         73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
                ||
Db         13 CC 12


RESULT 302
US-09-925-065A-925328/c
; Sequence 925328, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925328
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925328

  Query Match             6.2%;  Score 465.7;  DB 1;  Length 552;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  472;  Conservative   38;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
                |||||||||||||| ||::|||||||||||||||||||||||||||||||||:| ||||
Db        552 CTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTTCTGG 493

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
                |||||:|||||||||| |||||||||||||||||||||||||||||||||  |||||||
Db        492 ACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAATTCTCATTTGCCT 433

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
                || ||||| |||  ||||||:| |||||||||| ||||||:|||||||||||||||||
Db        432 TTAAAGATCCTTTGAACCCAACGTCTCAACTCACGTGGACTGTTTTACCCCAAGGGTTCA 373
```

```
Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
Db        372 AGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAACTCTCATACC 313

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
Db        312 TGGACACTCTTGTCCTTTGGTACAGAGATGATTTAATTTTAGCCACCTGTTAAGAAACCT 253

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
Db        252 TGTGCCATCAAGCCACCCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCYACA-GGTTT 194

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db        193 CCAGATGAAAGGCTCAGCTCTGCTCACAGCACGTTAAATACTTAGGGCTAAAATTATCCA 134

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
Db        133 AAGACACCAAGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCCTATCCTTATCCCA 74

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db          73 AAACCCTAAAGCAACTATGATGGTTCCTTGGCGTAACAGGTTTCTGCCGAATATGGATTC 14

Qy       3271 CC 3272
Db          13 CC 12
```

RESULT 303
US-09-925-065A-411946
; Sequence 411946, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 411946
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-411946

Query Match 6.2%; Score 465; DB 1; Length 568;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 523; Conservative 2; Mismatches 42; Indels 8; Gaps 3;

```
Qy       1580 CCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCA 1639
Db          1 CCTGACTCTGTTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAGTCAACTGCA 60

Qy       1640 GACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTA 1698
Db          61 GACATTAGAAAAAAACTTCAAAAGTCCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTA 120

Qy       1699 TTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGAC 1758
Db         121 TTGAACTTGGCAACCTTGGTTTTTT-TAATAGACAYGAGGAGGAGCAGGTGGAACGGGAC 179

Qy       1759 AAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGG 1818
Db         180 AAATGGGATTAAAAAAAAGGCCACTGCTTTAGTCATGGCCCTCAGGCAAGCAGACTTTGG 239

Qy       1819 AGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
Db         240 AGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 299

Qy       1879 CGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCC 1938
Db         300 TGGTCTACAAGGATACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTAGTCC 359

Qy       1939 ATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTT 1998
Db         360 ATGCCCCTTATGTCAAGAGAATCACTGGAAGGCCCACTGCCCCAGGGAATGAAGGTCCTC 419

Qy       1999 TGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGC 2058
Db         420 TGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGAGGCAAGC 479

Qy       2059 GCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
Db         480 GCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATAGAGGCTAA--- 536

Qy       2119 AGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAG 2153
Db         537 ---CTGTCTCCTGGACACTGGTGCGGCCTTCTCAG 568
```

RESULT 304
US-09-925-065A-411946
; Sequence 411946, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 411946
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-411946

Query Match 6.2%; Score 465; DB 1; Length 568;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 523; Conservative 2; Mismatches 42; Indels 8; Gaps 3;

```
Qy       1580 CCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCA 1639
Db          1 CCTGACTCTGTTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAGTCAACTGCA 60

Qy       1640 GACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTA 1698
Db          61 GACATTAGAAAAAAACTTCAAAAGTCCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTA 120

Qy       1699 TTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGAC 1758
```

```
                |||||||||||||||:| ||||||||| ||||||| |: |||||||||||| ||||| ||||
Db        121 TTGAACTTGGCAACCTTGGTTTTTT-TAATAGACAYGAGGAGGAGCAGGTGGAACGGGAC 179

Qy       1759 AAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGG 1818
                ||| |||||||||||||||||||| |||||||||||| |||||||||||  ||||||||
Db        180 AAATGGGATTAAAAAAAAGGCCACTGCTTTAGTCATGGCCCTCAGGCAAGCAGACTTTGG 239

Qy       1819 AGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 AGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 299

Qy       1879 CGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCC 1938
                ||||||||||| |||||||||||||||||||||||||||||||| ||||| | ||||
Db        300 TGGTCTACAAGGATACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCTGCCCCCTAGTCC 359

Qy       1939 ATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTT 1998
                |||||||||| ||||| |||||||||||||||||||||||| |  |||||| |
Db        360 ATGCCCCTTATGTCAAGAGAATCACTGGAAGGCCCACTGCCCCAGGGAATGAAGGTCCTC 419

Qy       1999 TGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGC 2058
                |||||||||||||||||||||||||||||||||||||||||||||||||| | |||||||
Db        420 TGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGAGGCAAGC 479

Qy       2059 GCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
                |||| |||||||||||||||||||||||| |||||||||||||||| ||||    |
Db        480 GCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATAGAGGCTAA--- 536

Qy       2119 AGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAG 2153
                  |||||||||||||||||||||| ||||| ||
Db        537 ---CTGTCTCCTGGACACTGGTGCGGCCTTCTCAG 568


RESULT 305
US-10-301-480-293356
; Sequence 293356, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 293356
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-293356

  Query Match             6.2%;  Score 464.8;  DB 1;  Length 599;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  518;  Conservative   29;  Mismatches   48;  Indels    6;  Gaps    4;

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
                |||||||||| |||||||||||||||||||||||:||||:||| |||||| ||||:|||
Db          1 TTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCCCTCA 60

Qy       4044 YTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
                :| |||||||||||||||||||| |:||||||||:|||||||||||||||||||:|:|||
Db         61 CTCGGGCACTAGAATTAGGAGAAGGAAAAAAGGGTAAATATATATACAGACTCTAAGTAT 120

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
                |||:|||||||| |||||||||||::|||||||:|:||||||||||||||||| |||
Db         121 GCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTC- 179

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
                |||:|||||||||||:|||||||||||||||||||||:|||||||:||||:||||||:||
Db         180 GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGRAA 239

Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
                ||| :||||||||:|||||||||||||:    |||||| |||||||||:||||||||||
Db         240 CCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGAAATA 299

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
                :|||:|||: |||||||||:|||||||| :|||||||||||||||||||||||||||||||
Db         300 GAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCCATTA 359

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
                |||||||||||  ||   ||| |||||| |||||||||  |||||||||||||| |||||
Db         360 GAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTCCCAG 419

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
                |||||||||||| ||| ||||| ||||||||||||||    | ||  |||||||  ||
Db         420 TACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTCCAGA 479

Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
                || |||||||   |||||| |||||| |||||||||||| ||| |||||  ||||||||||
Db         480 TGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATTACTT 538

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
                |||||||||| | |||||| |||||||||| ||||| ||||||||:|||||||||| |
Db         539 AAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAAATTA 598

Qy       4580 T 4580
                |
Db         599 T 599


RESULT 306
US-10-301-480-906765
; Sequence 906765, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 906765
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-906765

  Query Match             6.2%;  Score 464.8;  DB 1;  Length 599;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  518;  Conservative   29;  Mismatches   48;  Indels    6;  Gaps    4;

Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
                |||||||||| |||||||||||||||||||||||:||||:||| |||||| ||||:|||
Db          1 TTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCCCTCA 60

Qy       4044 YTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
                :| |||||||||||||||||||| |:||||||||:|||||||||||||||||||:|:|||
Db         61 CTCGGGCACTAGAATTAGGAGAAGGAAAAAAGGGTAAATATATATACAGACTCTAAGTAT 120

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
                |||:|||||||| |||||||||||::|||||||:|:||||||||||||||||| |||
Db         121 GCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTC- 179

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
                |||:|||||||||||:|||||||||||||||||||||:|||||||:||||:||||||:||
Db         180 GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGRAA 239
```

```
Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
             ||| :|||||||:|||||||||||||:    |||||| |||||||||:||||||||||
Db        240 CCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGAAATA 299

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
             :|||:|||: ||||||||:||||||||| :||||||||||||||||||||||||||||||
Db        300 GAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCCATTA 359

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
             |||||||||||  ||   ||| |||||| |||||||||  ||||||||||||||| |||||
Db        360 GAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTCCCAG 419

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
             |||||||||||| ||| ||||| ||||||||||||||    | ||  ||||||||  ||
Db        420 TACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTCCAGA 479

Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
              || |||||||   |||||| |||||| |||||||||| ||| ||||| ||||||||||
Db        480 TGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATTACTT 538

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
             |||||||||||  | |||||| |||||||||| ||||| ||||||||:|||||||||  |
Db        539 AAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAAATTA 598

Qy       4580 T 4580
              |
Db        599 T 599
```

```
RESULT 307
US-10-301-480-251972/c
; Sequence 251972, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251972
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-251972

  Query Match             6.2%;  Score 464.7;  DB 1;  Length 591;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  519;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||||||| ||||||||||:|||:|||||||||
Db        591 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCT 532

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAA 4691
             || ||: ||||||:|:||||||| ||||||||:|||||||||||||||||||:||||||
Db        531 GTAAAGAGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAA 473

Qy       4692 RAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTC 4750
             :|||||||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | |
Db        472 GAACAGGCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCC 413

Qy       4751 AGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCC 4810
             | |||||||||||||||||| |||||||:|| |||||||| |||||||||:|||||||||
Db        412 AAGGATTTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCC 353
```

```
Qy       4811 CCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCA 4870
             | ||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db        352 CTTGTAGGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCA 293

Qy       4871 GATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAA 4930
             ||||| |||||||||||||| |||||||||||||| |||| |||||| ||||  ||||||
Db        292 GATTCAGACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAA 233

Qy       4931 CCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCAC 4990
             |||||||||||||||||| ||||| |||| |||||||||||| |||| |||| |||||||
Db        232 CCCAGGGAGTATCCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCAC 173

Qy       4991 AGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAA 5050
             | |||||||| || |||:|||||||||||||||: |||| | |||||||||||||| ||
Db        172 AATCCTCAGGAAAAGTCRAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAA 113

Qy       5051 CCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCA 5110
             |||| |||||| |||| |||||||| :|||||||| |||||||||| |||||||| |
Db        112 CCCAAGAAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAA 53

Qy       5111 ACTTTCCCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
             ||| ||||||||||||||||||      ||| | | ||||||||||| |||
Db         52 ACTCTCCCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1
```

```
RESULT 308
US-10-301-480-251974/c
; Sequence 251974, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251974
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-251974

  Query Match             6.2%;  Score 464.7;  DB 1;  Length 591;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  519;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
             |||||||||||||||||||||||||||||||||||| ||||||||||:|||:|||||||||
Db        591 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCT 532

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAA 4691
             || ||: ||||||:|:||||||| ||||||||:|||||||||||||||||||:||||||
Db        531 GTAAAGAGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAA 473

Qy       4692 RAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTC 4750
             :|||||||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | |
Db        472 GAACAGGCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCC 413

Qy       4751 AGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCC 4810
             | |||||||||||||||||| |||||||:|| |||||||| |||||||||:|||||||||
Db        412 AAGGATTTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCC 353

Qy       4811 CCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCA 4870
             | ||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db        352 CTTGTAGGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCA 293

Qy       4871 GATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAA 4930
```

```
                |||||  |||||||||||||||  |||||||||||||||  ||||  ||||||  ||||  ||||||
Db        292 GATTCAGACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAA 233

Qy       4931 CCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCAC 4990
                |||||||||||:||||||  |||||  ||||  ||||||||||  ||||  ||||  |||||||
Db        232 CCCAGGGAGTATYCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCAC 173

Qy       4991 AGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAA 5050
                | ||||||||  ||  ||||||||||||||||||||: ||||  | ||||||||||||||  ||
Db        172 AATCCTCAGGAAAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAA 113

Qy       5051 CCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCA 5110
                ||||  ||||||  ||||  ||||||||  :||||||||  ||||||||||  |||||||||  |
Db        112 CCCAAGAAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAA 53

Qy       5111 ACTTTCCCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
                |||  ||||||||||||||||||      |||  |  |  ||||||||||||  |||
Db          52 ACTCTCCCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 309
US-10-301-480-865381/c
; Sequence 865381, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865381
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865381

  Query Match             6.2%;  Score 464.7;  DB 1;  Length 591;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  519;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCT 532

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAA 4691
                || ||: |||||||:|:|||||||  ||||||||:|||||||||||||||||||||:||||||
Db        531 GTAAAGAGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAA 473

Qy       4692 RAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTC 4750
                :|||||||||||||||  |:|:||:||||||||||  ||||||||  | ||  |||||||  | |
Db        472 GAACAGGCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCC 413

Qy       4751 AGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCC 4810
                | ||||||||||||||||||  |||||||:||  |||||||  |||||||||:|||||||||
Db        412 AAGGATTTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCC 353

Qy       4811 CCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCA 4870
                | |||||||  |||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        352 CTTGTAGGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCA 293

Qy       4871 GATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAA 4930
                |||||  ||||||||||||||  ||||||||||||||  ||||  ||||||  ||||  ||||||
Db        292 GATTCAGACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAA 233

Qy       4931 CCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCAC 4990
                |||||||||||||||||||  |||||  ||||  ||||||||||  ||||  ||||  |||||||
Db        232 CCCAGGGAGTATCCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCAC 173

Qy       4991 AGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAA 5050
                | ||||||||  ||  |||:|||||||||||||||||: ||||  | ||||||||||||||  ||
Db        172 AATCCTCAGGAAAAGTCRAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAA 113

Qy       5051 CCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCA 5110
                ||||  ||||||  ||||  ||||||||  :||||||||  ||||||||||  |||||||||  |
Db        112 CCCAAGAAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAA 53

Qy       5111 ACTTTCCCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
                |||  ||||||||||||||||||      |||  |  |  ||||||||||||  |||
Db          52 ACTCTCCCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 310
US-10-301-480-865383/c
; Sequence 865383, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865383
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865383

  Query Match             6.2%;  Score 464.7;  DB 1;  Length 591;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  519;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4572 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCT 4631
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 CCAAATCATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCT 532

Qy       4632 GTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAA 4691
                || ||: |||||||:|:|||||||  ||||||||:|||||||||||||||||||||:||||||
Db        531 GTAAAGAGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAA 473

Qy       4692 RAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTC 4750
                :|||||||||||||||  |:|:||:||||||||||  ||||||||  | ||  |||||||  | |
Db        472 GAACAGGCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCC 413

Qy       4751 AGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCC 4810
                | ||||||||||||||||||  |||||||:||  |||||||  |||||||||:|||||||||
Db        412 AAGGATTTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCC 353

Qy       4811 CCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCA 4870
                | |||||||  |||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        352 CTTGTAGGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCA 293

Qy       4871 GATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAA 4930
                |||||  ||||||||||||||  ||||||||||||||  ||||  ||||||  ||||  ||||||
Db        292 GATTCAGACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAA 233

Qy       4931 CCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCAC 4990
                |||||||||||:||||||  |||||  ||||  ||||||||||  ||||  ||||  |||||||
Db        232 CCCAGGGAGTATYCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCAC 173

Qy       4991 AGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAA 5050
                | ||||||||  ||  |||:|||||||||||||||||: ||||  | ||||||||||||||  ||
Db        172 AATCCTCAGGAAAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAA 113
```

```
Qy       5051 CCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCA 5110
Db        112 CCCAAGAAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAA 53

Qy       5111 ACTTTCCCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
Db         52 ACTCTCCCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1
```

RESULT 311
US-10-301-480-293355
; Sequence 293355, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 293355
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-293355

Query Match 6.2%; Score 463.6; DB 1; Length 599;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 519; Conservative 27; Mismatches 49; Indels 6; Gaps 4;

```
Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
Db          1 TTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCCCTCA 60

Qy       4044 YTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
Db         61 CTCGGGCACTAGAATTAGGAGAAGGRAAAAAGGGTAAATATATATACAGACTCTAAGTAT 120

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
Db        121 GCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTC- 179

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
Db        180 GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGGAA 239

Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
Db        240 CCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGAAATA 299

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
Db        300 GAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCCATTA 359

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
Db        360 GAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTCCCAG 419

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
Db        420 TACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTCCAGA 479

Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
Db        480 TGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATTACTT 538

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
Db        539 AAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAAATTA 598

Qy       4580 T 4580
Db        599 T 599
```

RESULT 312
US-10-301-480-906764
; Sequence 906764, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 906764
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-906764

Query Match 6.2%; Score 463.6; DB 1; Length 599;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 519; Conservative 27; Mismatches 49; Indels 6; Gaps 4;

```
Qy       3984 TTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCA 4043
Db          1 TTGAAAGTAAACCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAACAGCCCTCA 60

Qy       4044 YTKGGGCACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
Db         61 CTCGGGCACTAGAATTAGGAGAAGGRAAAAAGGGTAAATATATATACAGACTCTAAGTAT 120

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
Db        121 GCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTC- 179

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
Db        180 GAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGGAA 239

Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
Db        240 CCTGAAGAGGTGGCAGTCTTACACTGCCATCATCATCAGAAAGGAAAGGAAAGGGAAATA 299

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
Db        300 GAAGGGAACCGCCAAGCAGATATTGAAGTCAAAAGAGCTGCAAGGCAGGACCCTCCATTA 359

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
Db        360 GAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCACTCCGGGAAACCAAGTCCCAG 419

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
Db        420 TACTCAGCAGGAAAAATAGAATAGGGAACCTCACGAGGACATACTTCCCTCCCCTCCAGA 479

Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
Db        480 TGGCTAGCCACCAGAGAAGGAAAAATA-TTTTGCCTGCAGCTAACCAATAGAAATTACTT 538

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
```

```
                 |||||||||||| | |||||| |||||||||||| ||||| ||||||||:|||||||||| |
Db        539 AAAACCCTTCACCTAACCTTCCACTTAGGCATTGATAGTACCCATCAGATGGCCAAATTA 598

Qy       4580 T 4580
                 |
Db        599 T 599
```

RESULT 313
US-09-925-065A-758566
; Sequence 758566, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758566
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-758566

Query Match 6.2%; Score 461.4; DB 1; Length 633;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 520; Conservative 30; Mismatches 58; Indels 6; Gaps 4;

```
Qy       4050 CACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYAC 4108
                 |||||||||||||||| |:||||||| :|||||||||||||||||||:|:||||||:||
Db          5 CACTAGAATTAGGAGAAGGAAAAAAGGTTAAATATATATACAGACTCTAAGTATGCTTAC 64

Qy       4109 CTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGA 4168
                 ||||| |||||||||| |::|||||||:|:||||| |||||||||||||||:|||: |
Db         65 TTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAAAAGGAATTCCTAACTTCTGAGGAA 124

Qy       4169 ACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
                 ||||||||||:|||||||||||||||||||:| |||||:||| :||| ||||| |||:|
Db        125 ACACCTATCAAACATCAGGAAGCCATTAGGAGAATATTATTGGGTGTAGAGAAA-CTAAA 183

Qy       4229 GAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRG 4288
                 |||||||:|||||||||| |:|||||||||| |||||||| :|||| |||||:|||:|
Db        184 GAGGTGGCAGTCTTACACTTCCGGGGTCATCAGAAAGGAAAAGAAAGCTAAATAGAAGGG 243

Qy       4289 AAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATG 4348
                 ||:  | |||| :||||||||||:|||||||  |||||  |||||||||||||||||||
Db        244 AACCTCTAAGCGGATATTGAAGCCAAAAGAGCCACAAGGTGGGACCCTCCATTAGAAATG 303

Qy       4349 CTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCA 4408
                 |||||  ||    || |||||| ||:|||||| ||| ||||||||||||||||||||
Db        304 CTTATAGAAGGAACCCTAGTATGGGRTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCA 363

Qy       4409 GCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTA 4465
                 |||| ||||| |||||||||||||||| ||||    ||||| |||||||| ||| || ||
Db        364 GCAGAAGAAATAGAATGGGGAACCTCAAGAGGACATAGTTTCCTCCCCTCAGGATGGCTA 423

Qy       4466 GCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACC 4525
                 |||||  ||||||| |||||||||||||||||| ||| ||| |||||||||||||| |
Db        424 GCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCATTGGAAATTACTTAAAATC 483

Qy       4526 CTTCATCAAACCTTTCACTTAGGCA-TCGATAGCACCCATCARATGGCCAAATCATTATT 4584
                 ||||| || |||||||||||||||| | |||||||||||||:|||||||||||||||||
Db        484 CTTCACCACACCTTTCACTTAGGCATTTGATAGCACCCATCAGATGGCCAAATCATTATT 543

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
                 |||||||  ||| |||||||||||||||||| |:|||:|||||||||||||| |||| ||
Db        544 TACTGGATGAGGACTTTTCAAAACTATCAAGTAGATAGTCAGGGCCTGTGAAATGTGGCA 603

Qy       4645 RARAAATAATCCCC 4658
                 :|:||||| || ||
Db        604 AAGAAATATTCACC 617

RESULT 314
US-09-925-065A-758566
; Sequence 758566, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758566
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-758566

Query Match 6.2%; Score 461.4; DB 1; Length 633;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 520; Conservative 30; Mismatches 58; Indels 6; Gaps 4;

```
Qy       4050 CACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYAC 4108
                 |||||||||||||||| |:||||||| :|||||||||||||||||||:|:||||||:||
Db          5 CACTAGAATTAGGAGAAGGAAAAAAGGTTAAATATATATACAGACTCTAAGTATGCTTAC 64

Qy       4109 CTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGA 4168
                 ||||| |||||||||| |::|||||||:|:||||| |||||||||||||||:|||: |
Db         65 TTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAAAAGGAATTCCTAACTTCTGAGGAA 124

Qy       4169 ACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
                 ||||||||||:|||||||||||||||||||:| |||||:||| :||| ||||| |||:|
Db        125 ACACCTATCAAACATCAGGAAGCCATTAGGAGAATATTATTGGGTGTAGAGAAA-CTAAA 183

Qy       4229 GAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRG 4288
                 |||||||:|||||||||| |:|||||||||| |||||||| :|||| |||||:|||:|
Db        184 GAGGTGGCAGTCTTACACTTCCGGGGTCATCAGAAAGGAAAAGAAAGCTAAATAGAAGGG 243

Qy       4289 AAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATG 4348
                 ||:  | |||| :||||||||||:|||||||  |||||  |||||||||||||||||||
Db        244 AACCTCTAAGCGGATATTGAAGCCAAAAGAGCCACAAGGTGGGACCCTCCATTAGAAATG 303

Qy       4349 CTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCA 4408
                 |||||  ||    || |||||| ||:|||||| ||| ||||||||||||||||||||
Db        304 CTTATAGAAGGAACCCTAGTATGGGRTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCA 363

Qy       4409 GCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTA 4465
                 |||| ||||| |||||||||||||||| ||||    ||||| |||||||| ||| || ||
Db        364 GCAGAAGAAATAGAATGGGGAACCTCAAGAGGACATAGTTTCCTCCCCTCAGGATGGCTA 423
```
```

```
Qy       4466 GCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACC 4525
             |||||  ||||||| |||||||||||||||||| ||| ||| ||||||||||||||||| |
Db        424 GCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCATTGGAAATTACTTAAAATC 483

Qy       4526 CTTCATCAAACCTTTCACTTAGGCA-TCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             ||||| || |||||||||||||||| | |||||||||||||:||||||||||||||||||
Db        484 CTTCACCACACCTTTCACTTAGGCATTTGATAGCACCCATCAGATGGCCAAATCATTATT 543

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             ||||||| ||| |||||||||||||||| |:|||:||||||||||||| |||| ||
Db        544 TACTGGATGAGGACTTTTCAAAACTATCAAGTAGATAGTCAGGGCCTGTGAAATGTGGCA 603

Qy       4645 RARAAATAATCCCC 4658
              :|:||||| || ||
Db        604 AAGAAATATTCACC 617
```

RESULT 315
US-09-864-761-16537
; Sequence 16537, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16537
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007567.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
US-09-864-761-16537

Query Match 6.2%; Score 460.7; DB 1; Length 568;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 494; Conservative 28; Mismatches 31; Indels 3; Gaps 3;

```
Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||   |  | |||||||||| |||||||||| |||||||||||||||||||||||||
Db         16 ATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTAT 74

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              || | |||||||||||||| | |||||||||||||||||| |||||||||| ||::||||||
Db         75 CAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCC 133

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              |||||||| |||||||||||||||||| | ||||||||||:|||||||||| |||||||
Db        134 AAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATGCCTTTTTCTGCA 193

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              |||||||||||||||||||| |||||| |||||||||||||| |||| ||||||:| |
Db        194 TCCCTGTACATCCTGACTCTCGATTCTTTTTGCCTTTGAAGATCCTTCGAACCCAACGT 253

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
              |||||||||| ||||||:|||||||||||||||||||||||:|||||||||||||||
Db        254 CTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 313

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
              || |||||||||||||||||||:||:|::|||||||||||||||||||||||||:||| |
Db        314 AGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCAGTAAG 373

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
              || |||||||||||| :||:|||::|||||||||||||||||||||||||||||||:|
Db        374 TGAATGATTTACTTTCAGCTGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCAC 433

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
              ||||:|||||||| ||:|||||||||||||:||||||||||:| |:|||||:|||||||
Db        434 TCTTAAATTTCCTTGCCACCTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCT 492

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
              || | |||| ||||||||||||:||||:|||||||||||||||:|||||||||||||||
Db        493 CAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 552

Qy       3175 AAYRYATCCAGCCTAT 3190
              ||:::|||||||| ||
Db        553 AACGTATCCAGCCCAT 568
```

RESULT 316
US-09-925-065A-159043/c
; Sequence 159043, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159043
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159043

  Query Match             6.1%;  Score 458.7;  DB 1;  Length 585;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  513;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4578 CATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAK 4637
             |||||||||||||||||||||||||||||| |||||||||:|||:|||||||||||| ||:
Db        585 CATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCTGTAAAG 526

Qy       4638 TGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAG 4697
              ||||||:|:|||||||| |||||||:|||||||||||||||||:|||||||:|||||
Db        525 AGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAG 467

Qy       4698 GCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGAT 4756
              |||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | || ||||
Db        466 GCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCCAAGGAT 407

Qy       4757 TTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTA 4816
              |||||||||||||| |||||||:|| ||||||| |||||||||:|||||||||| ||||
Db        406 TTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTA 347

Qy       4817 GGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCG 4876
              ||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        346 GGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTCA 287

Qy       4877 GACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGG 4936
              |||||||||||||| |||||||||||||| |||| |||||| ||||  ||||||||||||
Db        286 GACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGG 227

Qy       4937 GAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCT 4996
              |||||||||||| ||||| |||| |||||||||| |||| |||| |||||||| ||||
Db        226 GAGTATCCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCACAATCCT 167

Qy       4997 CAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGG 5056
              |||| || |||:|||||||||||||: |||| | |||||||||||||| |||||| |
Db        166 CAGGAAAAGTCRAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAG 107

Qy       5057 AAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTC 5116
              ||||| |||| |||||||| :|||||||| |||||||||  |||||||||| |||| ||
Db        106 AAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTC 47

Qy       5117 CCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
              |||||||||||||||||     ||| | | |||||||||||| |||
Db         46 CCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 317
US-09-925-065A-159043/c
; Sequence 159043, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159043
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159043

  Query Match             6.1%;  Score 458.7;  DB 1;  Length 585;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  513;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4578 CATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAK 4637
             |||||||||||||||||||||||||||||| |||||||||:|||:|||||||||||| ||:
Db        585 CATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCTGTAAAG 526

Qy       4638 TGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAG 4697
              ||||||:|:|||||||| |||||||:|||||||||||||||||:|||||||:|||||
Db        525 AGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAG 467

Qy       4698 GCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGAT 4756
              |||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | || ||||
Db        466 GCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCCAAGGAT 407

Qy       4757 TTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTA 4816
              |||||||||||||| |||||||:|| ||||||| |||||||||:|||||||||| ||||
Db        406 TTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTA 347

Qy       4817 GGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCG 4876
              ||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        346 GGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTCA 287

Qy       4877 GACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGG 4936
              |||||||||||||| |||||||||||||| |||| |||||| ||||  ||||||||||||
Db        286 GACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGG 227

Qy       4937 GAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCT 4996
              |||||||||||| ||||| |||| |||||||||| |||| |||| |||||||| ||||
Db        226 GAGTATCCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCACAATCCT 167

Qy       4997 CAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGG 5056
              |||| || |||:|||||||||||||: |||| | |||||||||||||| |||||| |
Db        166 CAGGAAAAGTCRAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAG 107

Qy       5057 AAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTC 5116
              ||||| |||| |||||||| :|||||||| |||||||||  |||||||||| |||| ||
Db        106 AAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTC 47

Qy       5117 CCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
              |||||||||||||||||     ||| | | |||||||||||| |||
Db         46 CCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 318
US-09-925-065A-159045/c
; Sequence 159045, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159045
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159045

  Query Match             6.1%;  Score 458.7;  DB 1;  Length 585;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  513;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4578 CATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAK 4637
             |||||||||||||||||||||||||||||| ||||||||||:|||:|||||||||||| ||:
Db        585 CATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCTGTAAAG 526

Qy       4638 TGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAG 4697
              ||||||:|:||||||| ||||||||:|||||||||||||||||||:||||||:|||||
Db        525 AGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAG 467

Qy       4698 GCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGAT 4756
              |||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | || ||||
Db        466 GCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCCAAGGAT 407

Qy       4757 TTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTA 4816
              |||||||||||||| |||||||:|| ||||||| |||||||||:|||||||||| ||||
Db        406 TTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTA 347

Qy       4817 GGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCG 4876
              ||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        346 GGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTCA 287

Qy       4877 GACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGG 4936
              |||||||||||||| |||||||||||||| |||| |||||| ||||  ||||||||||||
Db        286 GACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGG 227

Qy       4937 GAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCT 4996
              ||||||:|||||| ||||| |||| |||||||||| |||| |||| |||||||| ||||
Db        226 GAGTATYCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCACAATCCT 167

Qy       4997 CAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGG 5056
              |||| || ||||||||||||||||||: |||| | |||||||||||||| |||||| |
Db        166 CAGGAAAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAG 107

Qy       5057 AAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTC 5116
              ||||| ||||  |||||||| :|||||||| |||||||||  |||||||| |||| ||
Db        106 AAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTC 47

Qy       5117 CCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
              |||||||||||||||||     ||| | | ||||||||||| |||
Db         46 CCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 319
US-09-925-065A-159045/c
; Sequence 159045, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159045
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159045

  Query Match             6.1%;  Score 458.7;  DB 1;  Length 585;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  513;  Conservative   16;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       4578 CATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAK 4637
             |||||||||||||||||||||||||||||| ||||||||||:|||:|||||||||||| ||:
Db        585 CATTATTTACTGGACCAGGCCTTTTCAAAGCTATCAAGCAGATAGTCAGGGCCTGTAAAG 526

Qy       4638 TGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAG 4697
              ||||||:|:||||||| ||||||||:|||||||||||||||||||:||||||:|||||
Db        525 AGTGCCAAAGAAATAAT-CCCTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAG 467

Qy       4698 GCCATTACCCTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGAT 4756
              |||||||||| |:|:||:|||||||||| |||||||| | || ||||||| | || ||||
Db        466 GCCATTACCCAGGAGAAGACTGGCAACTAGATTTTACTCGCATGCCCAAATCCCAAGGAT 407

Qy       4757 TTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTA 4816
              |||||||||||||| |||||||:|| ||||||| |||||||||:|||||||||| ||||
Db        406 TTCAGTATCTACTAGCCTGGGTAGATGCTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTA 347

Qy       4817 GGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCG 4876
              ||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        346 GGATAGAAAAGGCCCAAGAGGTAATAAAGGCACTAATTCATGAAATAATTCCCAGATTCA 287

Qy       4877 GACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGG 4936
              |||||||||||||| |||||||||||||| |||| |||||| ||||  ||||||||||||
Db        286 GACTTCCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGG 227

Qy       4937 GAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCT 4996
              ||||||:|||||| ||||| |||| |||||||||| |||| |||| |||||||| ||||
Db        226 GAGTATYCCAGGCCTTAGGCATACAATATCACTTACGCTGCACCTGGAGGCCACAATCCT 167

Qy       4997 CAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGG 5056
              |||| || ||||||||||||||||||: |||| | |||||||||||||| |||||| |
Db        166 CAGGAAAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAG 107

Qy       5057 AAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTC 5116
              ||||| ||||  |||||||| :|||||||| |||||||||  |||||||| |||| ||
Db        106 AAACCTACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTC 47

Qy       5117 CCCAAAAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGC 5161
              |||||||||||||||||     ||| | | ||||||||||| |||
Db         46 CCCAAAAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGC 1


RESULT 320
US-09-925-065A-295923
; Sequence 295923, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295923
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295923

  Query Match             6.1%;  Score 455.7;  DB 1;  Length 569;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  504;  Conservative   21;  Mismatches   44;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             ||||||||||||||||| ||||||||| ||||   ||||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| ||||||||||||||||||||||||||:||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATKCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  |||||||||  ||||||  ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |||||||||| ||||| |||||| | |||||||||||||||| ||||||||| ||
Db        173 GGATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
              ::|||||||||||||||||||||||||||||||||:|||||||||||:||||||||||
Db        232 CTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             ||||| |||||||||||| ||| |||||||||||||||||||||||| ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
             |||:|||||||||||||||||:|||||||||||||||||| ||||:|| ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             ||||||||||||||||||||||||||||||:||:|::||||||||| | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
             |:|||  |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 321
US-09-925-065A-295923
; Sequence 295923, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295923
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295923

  Query Match             6.1%;  Score 455.7;  DB 1;  Length 569;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  504;  Conservative   21;  Mismatches   44;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             ||||||||||||||||| ||||||||| ||||   ||||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| ||||||||||||||||||||||||||||:||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATKCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  |||||||||  ||||||  ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |||||||||| ||||| |||||| | |||||||||||||||| ||||||||| ||
Db        173 GGATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
              ::|||||||||||||||||||||||||||||||||:|||||||||||:||||||||||
Db        232 CTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             ||||| |||||||||||| ||| |||||||||||||||||||||||| ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
             |||:|||||||||||||||||:|||||||||||||||||| ||||:|| ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             ||||||||||||||||||||||||||||||:||:|::||||||||| | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
             |:|||  |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 322 -
US-09-925-065A-295924
; Sequence 295924, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295924
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295924

  Query Match             6.1%;  Score 454.1;  DB 1;  Length 569;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  503;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             ||||||||||||||| |||||||| ||||   ||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  ||||||||||  |||||| ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |:||||||||| |||||| |||||| | |||||||||||||||| ||||||||| ||
Db        173 GRATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
             ::||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db        232 CTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             ||||| |||||||||||||| ||| |||||||||||||||||||||||||||| ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
             |||:||||||||||||||||||:|||||||||||||||||||| ||||:|| ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             |||||||||||||||||||||||||||||:||:|::|||||||| | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
             |:|||  |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 323
US-09-925-065A-295924
; Sequence 295924, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295924
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295924

  Query Match             6.1%;  Score 454.1;  DB 1;  Length 569;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  503;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             ||||||||||||||| |||||||| ||||   ||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  ||||||||||  |||||| ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |:||||||||| |||||| |||||| | |||||||||||||||| ||||||||| ||
Db        173 GRATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
             ::||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db        232 CTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             ||||| |||||||||||||| ||| |||||||||||||||||||||||||||| ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
             |||:||||||||||||||||||:|||||||||||||||||||| ||||:|| ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             |||||||||||||||||||||||||||||:||:|::|||||||| | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
             |:|||  |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 324
US-09-925-065A-295925
; Sequence 295925, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
```

```
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295925
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295925

  Query Match             6.1%;  Score 454.1;  DB 1;  Length 569;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  503;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             |||||||||||||||| ||||||||| ||||   |||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  ||||||||||  ||||||  ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |||||||||||  ||||||  ||||||  | ||||||||||||||||||  ||||||||||  ||
Db        173 GGATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
              ::|||||||||||:|||||||||||||||||||||:||||||||||:||||||||||
Db        232 CTTTCCCAAATACYAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             |||||  |||||||||||||  ||| |||||||||||||||||||||||||  ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
              |||:|||||||||||||||||||:|||||||||||||||||||||  ||||:||  ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             |||||||||||||||||||||||||||||:||:|::|||||||||  | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
              |:|||  ||||||||||||||||:||:|||::|| ||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 325
US-09-925-065A-295925
; Sequence 295925, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295925
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295925

  Query Match             6.1%;  Score 454.1;  DB 1;  Length 569;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  503;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
             |||||||||||||||| ||||||||| ||||   |||||||||||||| ||||||||||
Db          1 TTTCCCTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATT 60

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
             ||||| ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 GTTAAAGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC- 119

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCA 2687
                 |||   |  | ||||||||||  ||||||||||  ||||||  ||||||||||
Db        120 ------AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCA 172

Qy       2688 GGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTG 2747
             |||||||||||  ||||||  ||||||  | ||||||||||||||||||  ||||||||||  ||
Db        173 GGATTATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTG 231

Qy       2748 MYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTC 2807
              ::|||||||||||:|||||||||||||||||||||:||||||||||:||||||||||
Db        232 CTTTCCCAAATACYAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTT 291

Qy       2808 TTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAAC 2867
             |||||  |||||||||||||  ||| |||||||||||||||||||||||||  ||||||||
Db        292 TTCTGTATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAAC 351

Qy       2868 CCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTA 2927
              |||:|||||||||||||||||||:|||||||||||||||||||||  ||||:||  ||||||
Db        352 CCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTA 411

Qy       2928 TTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             |||||||||||||||||||||||||||||:||:|::|||||||||  | ||||||||||||
Db        412 TTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTT 471

Qy       2988 CRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACC 3047
              |:|||  ||||||||||||||||:||:|||::|| ||||||||||||||||||||||||
Db        472 CGGTACCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACC 531

Qy       3048 CAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
             |||| :|||||:||||||||||:||||||||||||||:
Db        532 CAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 569


RESULT 326
US-09-864-761-13065/c
; Sequence 13065, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
```

```
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13065
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC021774.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99
US-09-864-761-13065

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| ||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||||| |||:|| :|||||:|||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:||||||||||||||||||||:|:||||||||||:||||| |||| ||||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||: |||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 327
US-09-925-065A-421237
; Sequence 421237, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421237
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-421237

  Query Match             6.1%;  Score 453.6;  DB 1;  Length 606;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  514;  Conservative   12;  Mismatches   61;  Indels   22;  Gaps    2;

Qy       6784 GAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATT 6843
             ||||||| ||||| |||||||||||||||| |||||||||| |||||||||:||||||||
Db          1 GAAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATT 60

Qy       6844 CSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTC 6903
             | ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||
Db         61 CTAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTC 120

Qy       6904 AGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
             ||||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 AGCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTA 180

Qy       6964 CTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAA 7023
             |||||||||||| ||||||||||:|||||||||||| ||||||||||||||||||||||
```

```
Db        181 CTCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAA 240

Qy       7024 GCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC 7071
              |||||:|| |||             |||||||| || ||| |||||||  || |||||||
Db        241 GCTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATC 300

Qy       7072 TACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGC 7131
              |||| | |||||||||||:||||||:||||||| | ||  ||||||||||||||||||||
Db        301 TACCACGGACCCCTGGACYGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGC 360

Qy       7132 ACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGC 7191
              ||||||||  ||||||||||  || |||||| ||    | ||||||  ||||||||||||
Db        361 ACCCCTCCCTAGGAAATCTCAAATGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGC 420

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
              |||||||| |||: || |||||||| |||||||||||| |||||||||||||||| |||
Db        421 AGTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGG 480

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
              | ||||||||||||||||||||||||||||||||  |:|||||||||          :|
Db        481 GCACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CT 530

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
              ||||||||||   |||||| |||||||||||| |||| |||||||:| |||   |||||
Db        531 GGGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCA 590

Qy       7372 ATCAGAGAG 7380
              ||||| ||
Db        591 ATCAGGTAG 599
```

RESULT 328
US-09-925-065A-421237
; Sequence 421237, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421237
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-421237

Query Match 6.1%; Score 453.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 514; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6784 GAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATT 6843
              ||||||| ||||| |||||||||||||||| |||||||||| |||||||||:||||||||
Db          1 GAAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATT 60

Qy       6844 CSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTC 6903
              | ||||| || ||||| ||:|||| |:||||||| |||||| ||||||||||||||||||
Db         61 CTAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTC 120

Qy       6904 AGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
              ||||:|||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        121 AGCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTA 180

Qy       6964 CTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAA 7023
              |||||||||||| ||||||||||:||||||||||| |||||||||||||||||||||||
Db        181 CTCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAA 240

Qy       7024 GCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC 7071
              |||||:|| |||             |||||||| || ||| |||||||  || |||||||
Db        241 GCTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATC 300

Qy       7072 TACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGC 7131
              |||| | |||||||||||:|||||||:||||||| | ||  ||||||||||||||||||||
Db        301 TACCACGGACCCCTGGACYGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGC 360

Qy       7132 ACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGC 7191
              ||||||||  ||||||||||  || |||||| ||    | ||||||  ||||||||||||
Db        361 ACCCCTCCCTAGGAAATCTCAAATGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGC 420

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
              |||||||| |||: || |||||||| |||||||||||| |||||||||||||||| |||
Db        421 AGTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGG 480

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
              | ||||||||||||||||||||||||||||||||  |:|||||||||          :|
Db        481 GCACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CT 530

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
              ||||||||||   |||||| |||||||||||| |||| |||||||:| |||   |||||
Db        531 GGGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCA 590

Qy       7372 ATCAGAGAG 7380
              ||||| ||
Db        591 ATCAGGTAG 599
```

RESULT 329
US-09-925-065A-421238
; Sequence 421238, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421238
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-421238

Query Match 6.1%; Score 453.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 514; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6784 GAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATT 6843
              ||||||| ||||| |||||||||||||||| |||||||||| |||||||||:||||||||
Db          1 GAAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATT 60
```

```
Qy       6844 CSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTC 6903
             | ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||
Db         61 CTAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTC 120

Qy       6904 AGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
             ||||:||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 AGCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTA 180

Qy       6964 CTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAA 7023
             |||||||||||| ||||||||||:|||||||||||| ||||||||||||||||||||||
Db        181 CTCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAA 240

Qy       7024 GCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC 7071
             |||||:|| |||             |||||||| || ||| |||||||  || |||||||
Db        241 GCTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATC 300

Qy       7072 TACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGC 7131
             |||| | |||||||||||| ||||||:||||||| | ||  ||||||||||||||||||||
Db        301 TACCACGGACCCCTGGACTGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGC 360

Qy       7132 ACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGC 7191
             ||||||||  |||||||||| :|| |||||| ||   | |||||| ||||||||||||||
Db        361 ACCCCTCCCTAGGAAATCTCAAMTGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGC 420

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
             |||||||| |||: || |||||||| |||||||||||||| ||||||||||||||||| |||
Db        421 AGTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGG 480

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
             | |||||||||||||||||||||||||||||||||  |:|||||||||          :|
Db        481 GCACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CT 530

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
             ||||||||||   |||||| |||||||||||| |||| ||||||||:| |||   |||||
Db        531 GGGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCA 590

Qy       7372 ATCAGAGAG 7380
             |||||  ||
Db        591 ATCAGGTAG 599
```

RESULT 330
US-09-925-065A-421238
; Sequence 421238, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421238
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-421238

Query Match 6.1%; Score 453.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 514; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6784 GAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATT 6843
             ||||||| ||||| |||||||||||||||| |||||||||| |||||||||:||||||||
Db          1 GAAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATT 60

Qy       6844 CSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTC 6903
             | ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||
Db         61 CTAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTC 120

Qy       6904 AGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
             ||||:||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 AGCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTA 180

Qy       6964 CTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAA 7023
             |||||||||||| ||||||||||:|||||||||||| ||||||||||||||||||||||
Db        181 CTCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAA 240

Qy       7024 GCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC 7071
             |||||:|| |||             |||||||| || ||| |||||||  || |||||||
Db        241 GCTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATC 300

Qy       7072 TACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGC 7131
             |||| | |||||||||||| ||||||:||||||| | ||  ||||||||||||||||||||
Db        301 TACCACGGACCCCTGGACTGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGC 360

Qy       7132 ACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGC 7191
             ||||||||  |||||||||| :|| |||||| ||   | |||||| ||||||||||||||
Db        361 ACCCCTCCCTAGGAAATCTCAAMTGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGC 420

Qy       7192 AGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGG 7251
             |||||||| |||: || |||||||| |||||||||||||| |||||||||||||||||| |||
Db        421 AGTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGG 480

Qy       7252 GGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGST 7311
             | |||||||||||||||||||||||||||||||||  |:|||||||||          :|
Db        481 GCACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CT 530

Qy       7312 GGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCA 7371
             ||||||||||   |||||| |||||||||||| |||| ||||||||:| |||   |||||
Db        531 GGGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCA 590

Qy       7372 ATCAGAGAG 7380
             |||||  ||
Db        591 ATCAGGTAG 599
```

RESULT 331
US-10-301-480-486080/c
; Sequence 486080, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 486080
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-486080

Query Match 6.0%; Score 452.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 513; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6785 AAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTC 6844
             |||||| ||||| |||||||||||||||| |||||||||| |||||||||:|||||||||
Db        606 AAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTC 547

Qy       6845 SAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCA 6904
             ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||||
Db        546 TAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTCA 487

Qy       6905 GCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTAC 6964
             |||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        486 GCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTAC 427

Qy       6965 TCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAG 7024
             |||||||||| ||||||||||:|||||||||||| ||||||||||||||||||||||||
Db        426 TCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAAG 367

Qy       7025 CTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCT 7072
             ||||:|| |||             |||||||| || ||| |||||||  || ||||||||
Db        366 CTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATCT 307

Qy       7073 ACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCA 7132
             ||| | |||||||||||:||||||:||||||| | ||  ||||||||||||||||||||||
Db        306 ACCACGGACCCCTGGACYGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGCA 247

Qy       7133 CCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCA 7192
             |||||||  |||||||||||  || |||||| ||    | ||||||  ||||||||||||||
Db        246 CCCCTCCCTAGGAAATCTCAAATGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGCA 187

Qy       7193 GTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGG 7252
             ||||||| |||: || |||||||| |||||||||||||| |||||||||||||||| ||||
Db        186 GTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGG 127

Qy       7253 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTG 7312
             |||||||||||||||||||||||||||||||||  |:|||||||||          :||
Db        126 CACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CTG 77

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
             |||||||||   |||||| |||||||||||| |||| ||||||||:| |||   ||||||
Db         76 GGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCAA 17

Qy       7373 TCAGAGAG 7380
             ||||  ||
Db         16 TCAGGTAG 9
```

RESULT 332
US-10-301-480-486081/c
; Sequence 486081, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 486081
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-486081

Query Match 6.0%; Score 452.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 513; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6785 AAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTC 6844
             |||||| ||||| |||||||||||||||| |||||||||| |||||||||:|||||||||
Db        606 AAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTC 547

Qy       6845 SAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCA 6904
             ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||||
Db        546 TAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTCA 487

Qy       6905 GCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTAC 6964
             |||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        486 GCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTAC 427

Qy       6965 TCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAG 7024
             |||||||||| ||||||||||:|||||||||||| ||||||||||||||||||||||||
Db        426 TCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAAG 367

Qy       7025 CTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCT 7072
             ||||:|| |||             |||||||| || ||| |||||||  || ||||||||
Db        366 CTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATCT 307

Qy       7073 ACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCA 7132
             ||| | |||||||||||:||||||:||||||| | ||  ||||||||||||||||||||||
Db        306 ACCACGGACCCCTGGACTGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGCA 247

Qy       7133 CCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCA 7192
             |||||||  |||||||||||  || |||||| ||    | ||||||  ||||||||||||||
Db        246 CCCCTCCCTAGGAAATCTCAAMTGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGCA 187

Qy       7193 GTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGG 7252
             ||||||| |||: || |||||||| |||||||||||||| |||||||||||||||| ||||
Db        186 GTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGG 127

Qy       7253 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTG 7312
             |||||||||||||||||||||||||||||||||  |:|||||||||          :||
Db        126 CACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CTG 77

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
             |||||||||   |||||| |||||||||||| |||| ||||||||:| |||   ||||||
Db         76 GGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCAA 17

Qy       7373 TCAGAGAG 7380
             ||||  ||
Db         16 TCAGGTAG 9
```

RESULT 333
US-10-301-480-1099489/c
; Sequence 1099489, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1099489
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1099489

Query Match 6.0%; Score 452.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 513; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6785 AAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTC 6844
             |||||| ||||| |||||||||||||||| |||||||||| |||||||||:|||||||||
```

```
Db        606 AAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTC 547

Qy       6845 SAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCA 6904
              ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||
Db        546 TAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTCA 487

Qy       6905 GCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTAC 6964
              |||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        486 GCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTAC 427

Qy       6965 TCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAG 7024
              |||||||||||| ||||||||||:|||||||||||||| ||||||||||||||||||||
Db        426 TCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAAG 367

Qy       7025 CTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCT 7072
              ||||:|| |||            ||||||||| || ||| |||||||  || ||||||||
Db        366 CTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATCT 307

Qy       7073 ACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCA 7132
              ||| | ||||||||||:||||||:||||||| | ||  ||||||||||||||||||||
Db        306 ACCACGGACCCCTGGACYGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGCA 247

Qy       7133 CCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCA 7192
              |||||||  ||||||||||  || |||||| ||   | |||||| ||||||||||||||
Db        246 CCCCTCCCTAGGAAATCTCAAATGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGCA 187

Qy       7193 GTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGG 7252
              ||||||| |||: || |||||||| |||||||||||||| ||||||||||||||| ||||
Db        186 GTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGG 127

Qy       7253 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTG 7312
              |||||||||||||||||||||||||||||||||  |:|||||||||          :||
Db        126 CACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CTG 77

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
              |||||||||   |||||| |||||||||||| |||| ||||||||:| |||   ||||||
Db          76 GGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCAA 17

Qy       7373 TCAGAGAG 7380
              ||||  ||
Db          16 TCAGGTAG 9
```

RESULT 334
US-10-301-480-1099490/c
; Sequence 1099490, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1099490
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1099490

Query Match 6.0%; Score 452.6; DB 1; Length 606;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 513; Conservative 12; Mismatches 61; Indels 22; Gaps 2;

```
Qy       6785 AAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTC 6844
              |||||| ||||| |||||||||||||||| |||||||||| |||||||||:|||||||||
Db        606 AAGAATTCTGTTGTTATGTTAATCAATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTC 547

Qy       6845 SAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCA 6904
              ||||| || ||||| ||:|||| |:||||||| ||||||  ||||||||||||||||||
Db        546 TAGATCAAACACAACATAGAGCACAGGAGCTTCAAAACACCAGACCCTGGGGCCTCCTCA 487

Qy       6905 GCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTAC 6964
              |||:|||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        486 GCCAATGGATGCCCTGGATTCTCTCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTAC 427

Qy       6965 TCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAG 7024
              |||||||||||| ||||||||||:|||||||||||||| ||||||||||||||||||||
Db        426 TCCTCTTTGGATCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCGAAG 367

Qy       7025 CTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCT 7072
              ||||:|| |||            ||||||||| || ||| |||||||  || ||||||||
Db        366 CTGTAAAGCTACAAATGGTTCCTCAAATGGAGTCCCAGAGGCAGTCCGTGATTAAGATCT 307

Qy       7073 ACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCA 7132
              ||| | ||||||||||| |||||:||||||| | ||  ||||||||||||||||||||
Db        306 ACCACGGACCCCTGGACTGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCAAAGGCA 247

Qy       7133 CCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCA 7192
              |||||||  |||||||||| :|| |||||| ||   | |||||| ||||||||||||||
Db        246 CCCCTCCCTAGGAAATCTCAAMTGTACAACCCCTGGCATGCCCCAGTTCAGCAGGAAGCA 187

Qy       7193 GTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGG 7252
              ||||||| |||: || |||||||| |||||||||||||| ||||||||||||||| ||||
Db        186 GTTAGAGTGGTCATCAGCCAACCTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGG 127

Qy       7253 GACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTG 7312
              |||||||||||||||||||||||||||||||||  |:|||||||||          :||
Db        126 CACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAATTAAGAATCC----------CTG 77

Qy       7313 GGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA 7372
              |||||||||   |||||| |||||||||||| |||| ||||||||:| |||   ||||||
Db          76 GGAAGGTGATTGCATCCATCTTTAAACACGGAGCTTCCAACTTAGCTTACAATCGACCAA 17

Qy       7373 TCAGAGAG 7380
              ||||  ||
Db          16 TCAGGTAG 9
```

RESULT 335
US-10-301-480-416983/c
; Sequence 416983, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 416983
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-416983

Query Match 6.0%; Score 451.8; DB 1; Length 564;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 498; Conservative 11; Mismatches 53; Indels 2; Gaps 2;

```
Qy       4661 CCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGG 4720
              |||:|||||||||||||||||:|||||:|||||||| ||||| |:|:||:|||||
Db        564 CCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAGGAGAAGACTGG 505
```

```
Qy       4721 CAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTA 4779
             ||||| |||||||||||||| ||||||| |||||||||||||||||||||||| ||||||
Db        504 CAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTA 445

Qy       4780 RATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTA 4839
             :|||||||||| ||||||||:|||||||||| |||||||||||||||| ||||||||||
Db        444 GATACTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTAGGACAGAAAAGTCCCAAGAGGTA 385

Qy       4840 ATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGT 4899
             |||||||||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        384 ATAAAGGCACTAATTCGTGAAATAATTCCCAGATTCGTACTTCCCCGAGGCTTACAGAGT 325

Qy       4900 GACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATA 4959
             |||||| |||| |||||| |||| |||||||||||||||||||||    ||||| |||
Db        324 GACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGAAATTAGGCATA 265

Qy       4960 CGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAAT 5019
             | |||||||||| ||||| |||| |||||||| |||||||| || || |||||||| ||
Db        264 CAATATCACTTATACTGCACCTGGAGGCCACAATCCTCAGGAAAAGTTGAGAAAATAAAC 205

Qy       5020 GAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYT 5079
              |||:||||||   |||||||||| || |||||| :|||||||||  ||  |||||:|
Db        204 CAAACACTCAAGCAACATCTAAAAATGCTAACCCAARAAACCCACCTTGCACAGCCTGCT 145

Qy       5080 CTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCA 5139
             ||||||||||||||||   |||||||   |||| |||||| ||||||||||||||||||
Db        144 CTGTTGCCTATAGCCTTACTAAGAATCCAAAACTATCCCCATAAAGCAGGACTTAGCCCA 85

Qy       5140 TACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACA 5198
             ||| |||||||||||||| |||||||| |||||||||||||||| ||||||| | |||||
Db         84 TACAAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTTCTTGACCGAGAGACA 25

Qy       5199 GCCAACTTAGTTGCAGACATCACC 5222
             |||||||||||||||||||||||||
Db         24 GCCAACTTAGTTGCAGACATCACC 1
```

RESULT 336
US-10-301-480-1030392/c
; Sequence 1030392, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1030392
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1030392

Query Match 6.0%; Score 451.8; DB 1; Length 564;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 498; Conservative 11; Mismatches 53; Indels 2; Gaps 2;

```
Qy       4661 CCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGG 4720
             |||:|||||||||||||||||:||||||:||||||||| ||||| |:|:||:|||||
Db         564 CCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAGGAGAAGACTGG 505

Qy       4721 CAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTA 4779
             ||||| |||||||||||||| ||||||| |||||||||||||||||||||||| ||||||
Db        504 CAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTA 445

Qy       4780 RATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTA 4839
             :|||||||||| ||||||||:|||||||||| |||||||||||||||| ||||||||||
Db        444 GATACTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTAGGACAGAAAAGTCCCAAGAGGTA 385

Qy       4840 ATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGT 4899
             |||||||||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        384 ATAAAGGCACTAATTCGTGAAATAATTCCCAGATTCGTACTTCCCCGAGGCTTACAGAGT 325

Qy       4900 GACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATA 4959
             |||||| |||| |||||| |||| |||||||||||||||||||||    ||||| |||
Db        324 GACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGAAATTAGGCATA 265

Qy       4960 CGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAAT 5019
             | |||||||||| ||||| |||| |||||||| |||||||| || || |||||||| ||
Db        264 CAATATCACTTATACTGCACCTGGAGGCCACAATCCTCAGGAAAAGTTGAGAAAATAAAC 205

Qy       5020 GAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYT 5079
              |||:||||||   |||||||||| || |||||| :|||||||||  ||  |||||:|
Db        204 CAAACACTCAAGCAACATCTAAAAATGCTAACCCAARAAACCCACCTTGCACAGCCTGCT 145

Qy       5080 CTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCA 5139
             ||||||||||||||||   |||||||   |||| |||||| ||||||||||||||||||
Db        144 CTGTTGCCTATAGCCTTACTAAGAATCCAAAACTATCCCCATAAAGCAGGACTTAGCCCA 85

Qy       5140 TACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACA 5198
             ||| |||||||||||||| |||||||| |||||||||||||||| ||||||| | |||||
Db         84 TACAAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTTCTTGACCGAGAGACA 25

Qy       5199 GCCAACTTAGTTGCAGACATCACC 5222
             |||||||||||||||||||||||||
Db         24 GCCAACTTAGTTGCAGACATCACC 1
```

RESULT 337
US-10-301-480-372764
; Sequence 372764, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 372764
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-372764

Query Match 6.0%; Score 450.7; DB 1; Length 564;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 499; Conservative 21; Mismatches 44; Indels 9; Gaps 3;

```
Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
             |||||||||||| |||||||||| ||||   |||||||||||||| ||||||||||||||||
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
              ||| ||||||||||||||||||||||||:||  |  ||||||| || | |
Db         61 AGACCTAAAAGCCCAAGGCTTAGTAAAACCATKCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
              |||   |  | ||||||||||  ||||||||||  ||||||  |||||||||||||||
Db        115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGGATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
             |||||| |||||| |||||| | |||||||||||||||||| |||||||||| ||::|||
```

```
Db        173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||
Db        232 CCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
             ||||||||||||| ||| |||||||||||||||||||||||||| ||||||||||:|
Db        292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
             ||||||||||||||||||:||||||||||||||||||| ||||:|| ||||||||||
Db        352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
             |||||||||||||||||||||||||:||:|::|||||||||| | ||||||||||||:|||
Db        412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471

Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
                |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||||||
Db        472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
              :|||||:||||||||||:|||||||||||||:
Db        532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 338
US-10-301-480-986173
; Sequence 986173, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 986173
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-986173

  Query Match             6.0%;  Score 450.7;  DB 1;  Length 564;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  499;  Conservative   21;  Mismatches   44;  Indels    9;  Gaps    3;

Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
             |||||||||| ||||||||| ||||   |||||||||||||| ||||||||||||||||
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
              ||| |||||||||||||||||||||||:||  |  ||||||| || | |
Db         61 AGACCTAAAAGCCCAAGGCTTAGTAAAACCATKCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
               |||   |  | ||||||||||  ||||||||||  |||||| ||||||||||||||||
Db        115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGGATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
             |||||| |||||| |||||| | |||||||||||||||||| |||||||||| ||::|||
Db        173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||
Db        232 CCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
             ||||||||||||| ||| |||||||||||||||||||||||||| ||||||||||:|
Db        292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
             ||||||||||||||||||:|||||||||||||||||||| ||||:|| ||||||||||
Db        352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
             |||||||||||||||||||||||||:||:|::|||||||||| | ||||||||||||:|||
Db        412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471

Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
                |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||||||
Db        472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
              :|||||:||||||||||:|||||||||||||:
Db        532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 339
US-09-925-065A-950020
; Sequence 950020, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 950020
;   LENGTH: 608
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-950020

  Query Match             6.0%;  Score 450.4;  DB 1;  Length 608;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  528;  Conservative    3;  Mismatches   60;  Indels    8;  Gaps    4;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             ||||||||||||||||:||| ||||||| ||||||||||||||||||||| ||||||
Db         12 AAAAAACTTCAAAAGTCYGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 71

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             | |||:| |||||||||||||||||||||| |||||||||||| |||| ||| |||
Db         72 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 130

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||| |||||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db        131 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 189

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             || ||||||| |||||||||||   |||||||||||||||||||||| ||||| |||||
Db        190 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 249

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
```

```
Db        250 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 309

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || | ||||||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        310 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 369

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        370 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGCGCCAGCCCA 429

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
              |||||||||||||||||||| |||||||||||||||||||||||||||||||    |  ||
Db        430 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGTTTAACTG 489

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
              || || ||||||||| |||| |||||||||||||| | ||| |||||||| ||||||||
Db        490 TCCCCAGGACACTGGCGCGGCCTTCTTAGTCTTACTGTCCTGCCCCGGACATCTGTCCTC 549

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCC 2243
              |||||||||||||| |     ||| | || ||| |||| |||||||||||||| : ||
Db        550 CAGATCTGTCACTATCC---AAGGGGTCCTAGGACAGGCAATCACTAGATACTTCTCCC 605


RESULT 340
US-09-925-065A-950020
; Sequence 950020, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 950020
;   LENGTH: 608
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-950020

  Query Match             6.0%;  Score 450.4;  DB 1;  Length 608;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  528;  Conservative    3;  Mismatches   60;  Indels    8;  Gaps    4;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              ||||||||||||||||:||| |||||||| |||||||||||||||||||||||| ||||||
Db         12 AAAAAACTTCAAAAGTCYGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 71

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              | |||:| |||||||||||||||||||||||||| |||||||||||| |||| ||| |||
Db         72 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 130

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              |||||| |||||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db        131 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 189

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              || ||||||| |||||||||||   |||||||||||||||||||||| ||||| |||||
Db        190 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 249

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        250 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 309

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || | ||||||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        310 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 369

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        370 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGCGCCAGCCCA 429

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TG 2124
              |||||||||||||||||||| |||||||||||||||||||||||||||||||    |  ||
Db        430 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGTTTAACTG 489

Qy       2125 TCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTC 2184
              || || ||||||||| |||| |||||||||||||| | ||| |||||||| ||||||||
Db        490 TCCCCAGGACACTGGCGCGGCCTTCTTAGTCTTACTGTCCTGCCCCGGACATCTGTCCTC 549

Qy       2185 CAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCC 2243
              |||||||||||||| |     ||| | || ||| |||| |||||||||||||| : ||
Db        550 CAGATCTGTCACTATCC---AAGGGGTCCTAGGACAGGCAATCACTAGATACTTCTCCC 605


RESULT 341
US-10-301-480-372765
; Sequence 372765, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 372765
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-372765

  Query Match             6.0%;  Score 449.1;  DB 1;  Length 564;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  498;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
              |||||||||| ||||||||| ||||   |||||||||||||| ||||||||||||||||
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
              ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 AGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
              |||    |  | |||||||||||  ||||||||||  |||||| ||||||||||:|||
Db         115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGRATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
              |||||| |||||| |||||| | |||||||||||||||| |||||||||| ||::|||
Db         173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
              |||||||||||||||||||||||||||||||:|||||||||||:|||||||||| |||||
Db         232 CCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
              |||||||||||||| ||| |||||||||||||||||||||||||| ||||||||||:|
```

```
Db        292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
             ||||||||||||||||||||:||||||||||||||||||||| ||||:|| ||||||||||
Db        352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
             ||||||||||||||||||||||:||:|::|||||||| | ||||||||||||:|||
Db        412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471

Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
               |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||||||
Db        472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
              :|||||:||||||||||:|||||||||||||:
Db        532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 342
US-10-301-480-372766
; Sequence 372766, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 372766
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-372766

  Query Match             6.0%;  Score 449.1;  DB 1;  Length 564;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  498;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
             |||||||||||| ||||||||| ||||   |||||||||||||| ||||||||||||||||
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
              ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 AGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
              |||   |  | ||||||||||  ||||||||||  ||||||  ||||||||||||||||
Db        115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGGATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
              ||||||  ||||||  ||||||  | ||||||||||||||||||  ||||||||||  ||::|||
Db        173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
              ||||||||:||||||||||||||||||||:||||||||||:||||||||||  |||||
Db        232 CCAAATACYAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
              |||||||||||||  ||| ||||||||||||||||||||||||||||  ||||||||||:|
Db        292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
              ||||||||||||||||||:||||||||||||||||||||| ||||:|| ||||||||||
Db        352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
              ||||||||||||||||||||||||:||:|::|||||||| | ||||||||||||:|||
Db        412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471

Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
               |||||||||||||||:||:|||::|| ||||||||||||||||||||||||||||||
Db        472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
              :|||||:||||||||||:|||||||||||||:
Db        532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 343
US-10-301-480-986174
; Sequence 986174, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 986174
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-986174

  Query Match             6.0%;  Score 449.1;  DB 1;  Length 564;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  498;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
             |||||||||||| ||||||||| ||||   |||||||||||||| ||||||||||||||||
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
              ||| |||||||||||||||||||||||||| ||  |  ||||||| || | |
Db         61 AGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
              |||   |  | ||||||||||  ||||||||||  ||||||  ||||||||||:|||
Db        115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGRATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
              ||||||  ||||||  ||||||  | ||||||||||||||||||  ||||||||||  ||::|||
Db        173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
              ||||||||||||||||||||||||||||||||||:||||||||||:||||||||||  |||||
Db        232 CCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
              |||||||||||||  ||| ||||||||||||||||||||||||||||  ||||||||||:|
Db        292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
              ||||||||||||||||||:||:|::|||||||| | ||||||||||||:|||
Db        352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
              ||||||||||||||||||||||||:||:|::|||||||| | ||||||||||||:|||
Db        412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471
```

```
Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
Db        472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
Db        532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 344
US-10-301-480-986175
; Sequence 986175, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 986175
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-986175

  Query Match             6.0%;  Score 449.1;  DB 1;  Length 564;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  498;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    3;

Qy       2513 CTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAA 2572
Db          1 CTACCAAAGGCGGTACCCCCTTAGACTGGAGGCCCAACAAGGACTCCAAAAGATTGTTAA 60

Qy       2573 GGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGT 2632
Db            AGACCTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGCCCCTGCAATACTCC------ 114

Qy       2633 GGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATT 2692
Db         115 -AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-AGTTAGTACAAGATCTCAGGATT 172

Qy       2693 ATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTC 2752
Db         173 ATCAAT-GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTC 231

Qy       2753 CCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTG 2812
Db         232 CCAAATACYAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTG 291

Qy       2813 CATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARC 2872
Db         292 TATCCCTGTACATCGTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAAC 351

Qy       2873 ATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGG 2932
Db         352 ATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGAATAGCCCTCATCTATTTGG 411

Qy       2933 CCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTA 2992
Db         412 CCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGAATACTCTTGTCCTTCGGTA 471

Qy       2993 KGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGC 3052
Db         472 CCTGGATGATTTACTTTTAGCTGCCCATTTAGAAACCTTGTGCCATCAAGCCACCCAAGT 531

Qy       3053 RCTCTTMAATTTCCTCGCYACCTGTGGCTACAW 3085
Db         532 GCTCTTAAATTTCCTCGCCACCTGTGGCTACAA 564


RESULT 345
US-09-925-065A-344446/c
; Sequence 344446, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344446
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-344446

  Query Match             6.0%;  Score 448.8;  DB 1;  Length 561;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  495;  Conservative   11;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       4664 YATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAA 4723
Db         561 TATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAGGAGAAGACTGGCAA 502

Qy       4724 CT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARAT 4782
Db         501 CTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGAT 442

Qy       4783 ACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATA 4842
Db         441 ACTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTAGGACAGAAAAGTCCCAAGAGGTAATA 382

Qy       4843 AAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGAC 4902
Db         381 AAGGCACTAATTCGTGAAATAATTCCCAGATTCGTACTTCCCCGAGGCTTACAGAGTGAC 322

Qy       4903 AATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGA 4962
Db         321 AATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGAAATTAGGCATACAA 262

Qy       4963 TATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAA 5022
Db         261 TATCACTTATACTGCACCTGGAGGCCACAATCCTCAGGAAAAGTTGAGAAAATAAACCAA 202

Qy       5023 AYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTG 5082
Db         201 ACACTCAAGCAACATCTAAAAATGCTAACCCAARAAACCCACCTTGCACAGCCTGCTCTG 142

Qy       5083 TTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATAC 5142
Db         141 TTGCCTATAGCCTTACTAAGAATCCAAAACTATCCCCATAAAGCAGGACTTAGCCCATAC 82

Qy       5143 GAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCC 5201
Db          81 AAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTTCTTGACCGAGAGACAGCC 22
```

```
Qy       5202 AACTTAGTTGCAGACATCACC 5222
             |||||||||||||||||||||
Db         21 AACTTAGTTGCAGACATCACC 1
```

RESULT 346
US-09-925-065A-344446/c
; Sequence 344446, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344446
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-344446

Query Match 6.0%; Score 448.8; DB 1; Length 561;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 495; Conservative 11; Mismatches 53; Indels 2; Gaps 2;

```
Qy       4664 YATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAA 4723
              :|||||||||||||||||||||:||||||:||||||||| ||||| |:|:||:||||||||
Db        561 TATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAGGAGAAGACTGGCAA 502

Qy       4724 CT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARAT 4782
              || |||||||||||||| ||||||| |||||||||||||||||||||||| |||||||:||
Db        501 CTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGAT 442

Qy       4783 ACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATA 4842
              |||||||| |||||||||:|||||||||| |||||||||||||| ||||||||||||||||
Db        441 ACTTTCACTGGTTGGGCAGAGGCCTTCCCTTGTAGGACAGAAAAGTCCCAAGAGGTAATA 382

Qy       4843 AAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGAC 4902
              ||||||||| ||| |||||||||||||||||||| |||||||||||||||||||||||||
Db        381 AAGGCACTAATTCGTGAAATAATTCCCAGATTCGTACTTCCCCGAGGCTTACAGAGTGAC 322

Qy       4903 AATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGA 4962
              ||| |||| |||||| |||| |||||||||||||||||||||||| ||||| |||| |
Db        321 AATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGAAATTAGGCATACAA 262

Qy       4963 TATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAA 5022
              ||||||||| ||||| |||| |||||||| |||||||| || || |||||||| ||  ||
Db        261 TATCACTTATACTGCACCTGGAGGCCACAATCCTCAGGAAAAGTTGAGAAAATAAACCAA 202

Qy       5023 AYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTG 5082
              |:|||||   |||||||||| || ||||||| :|||||||||  ||  |||||:||||
Db        201 ACACTCAAGCAACATCTAAAAATGCTAACCCAARAAACCCACCTTGCACAGCCTGCTCTG 142

Qy       5083 TTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATAC 5142
              ||||||||||||||   |||||||    |||| |||||| |||||||||||||||||||||
Db        141 TTGCCTATAGCCTTACTAAGAATCCAAAACTATCCCCATAAAGCAGGACTTAGCCCATAC 82

Qy       5143 GAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCC 5201
              |||||||||||||| |||||||| |||||||||||||||| ||||||| | ||||||||
Db         81 AAAATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTTCTTGACCGAGAGACAGCC 22
```

```
Qy       5202 AACTTAGTTGCAGACATCACC 5222
             |||||||||||||||||||||
Db         21 AACTTAGTTGCAGACATCACC 1
```

RESULT 347
US-09-925-065A-790020/c
; Sequence 790020, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790020
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-790020

Query Match 6.0%; Score 448.8; DB 1; Length 565;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 492; Conservative 42; Mismatches 30; Indels 20; Gaps 4;

```
Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
              |||||||||||||||||||||||||||||||||||||||||          |||:  ||||||||||
Db        565 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 515

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
              |||||||:|||||||||||||||||||| ||||||:|||| ||||||||||||||||||||
Db        514 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 455

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
              ||||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| |||||||||
Db        454 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 395

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
              | || |:||: ||::|||||||||||||| ||||||||||||||||:|||||:|||||
Db        394 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 335

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
              ||||||:|||||||||          ||||||| |:|:|||||:|||||||||||||||||
Db        334 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 284

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
              ||||||||||||:||||:|||||||||||| |||||:|||||||||| |||||||::|||||
Db        283 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 224

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
              || ||||||||||||||||||||:|||||||||||||||||||||:|::|||||||||:|
Db        223 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 164

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
              ||:|| :|||||||||:|||||||| ||||| |:  | :|||||||||||:|||||: ||
Db        163 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 105

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
              || |||||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::
```

```
Db        104 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 45

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
              ||||||||||||||| |||||||||||||| ||| || || :||
Db         44 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAACCCC 2


RESULT 348
US-09-925-065A-790020/c
; Sequence 790020, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790020
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-790020

  Query Match             6.0%;  Score 448.8;  DB 1;  Length 565;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  492;  Conservative   42;  Mismatches   30;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
              ||||||||||||||||||||||||||||||||||||||||        |||:  ||||||||||||
Db        565 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 515

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
              |||||||:|||||||||||||||||||| ||||||:|||| ||||||||||||||||||||||
Db        514 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 455

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
              ||||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| |||||||||
Db        454 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 395

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
              | || |:||: ||::|||||||||||||| ||||||||||||||||:|||||:|||||
Db        394 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 335

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
              ||||||:|||||||||         ||||||| |:|:|||||:|||||||||||||||||
Db        334 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 284

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
              ||||||||||||:||||:|||||||||||| |||||:|||||||| ||||||||::|||||
Db        283 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 224

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
              || |||||||||||||||||:|||||||||||||||||||:|::|||||||||:|
Db        223 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 164

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
              ||:|| :||||||||||:||||||| ||||| |:  | :||||||||||:|||||: ||
Db        163 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 105

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 45

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCC 3408
              ||||||||||||||| |||||||||||||| ||| || || :||
Db         44 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAACCCC 2


RESULT 349
US-10-029-386-10274
; Sequence 10274, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10274
;   LENGTH: 578
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC008626.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.82
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P26808, EVALUE 6.00e-31
;   OTHER INFORMATION: EST_HUMAN HIT: BF996761.1, EVALUE 0.00e+00
US-10-029-386-10274

  Query Match             6.0%;  Score 448.6;  DB 1;  Length 578;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  489;  Conservative   47;  Mismatches   42;  Indels    8;  Gaps    4;

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
              || |||||||: || :||| || ||||:|| ||:||||||:    |||| ||||||||||||
Db          1 TTTCCCAGGTACGGTGAAACAGTCAGGCCAGTATATACAC----TAAGTAAACTCAGAAA 56

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
              ||||||||||||||:|||||||||:|:|||||||:::||| |||||||||||||  ||
Db         57 GCCAATACCCATTTAGTAAGATGGACACCTGAAGCAAAAGAGGCTTTCCAGGCCCTAAAG 116

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||||:|||| ||||||||:|||||||:|||||||| ||| |:||:||
Db        117 AAGGCCCT--AACCCAAGCCCCAATGTTAAGTTTGCCAACAGGGCAAGGCTTCTCTTTAT 174

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |::|||||||||| |||||||:||:||||||:||| |||||||||||:| |:|||||: |||
Db        175 ATGTCACAGAAA-AAACAGGAATAGCTCTAGGAATCCTTACACAGGTTCAAGGGACCAGC 233

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
              |||||||::||||::|| |||:|||| |||:|||||||||||||||||||:|:||||
Db        234 TTGCAACCTGTGGCATACTTGAGTAAGAAAATTGATGTAGTGGCAAAGGGTTGGCCTCAT 293

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
              ||||||: |||||||  ||||||||||||::|||||||||||||||||||||||||||||||
Db        294 TGTTTACAGGTAGTGACGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAG 353

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:|||||||||||||||||||||||||||||:|| ||::| |||||||||||||||||||
Db        354 GGAAGAGATCTTACTGTGTGGACATCTCATGATGTAAATGGTATACTCACTGCTAAAGGA 413

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||: ||| |||||| :|||||||||||| |||||:| | |
Db        414 AACTTGTGGCTGTCAGACAACCATTTGCTTAAAAATCAGGCTCTATTAATTGAAGGACTA 473
```

```
Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
Db        474 GTGCTGAGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 533

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATG 3853
Db        534 GAAAAGAGAGAACGTAACTGTCAAC-GGTGATTGCTCAAACCTATG 578


RESULT 350
US-09-925-065A-787519/c
; Sequence 787519, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787519
;   LENGTH: 581
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-787519

  Query Match             6.0%;  Score 448.2;  DB 1;  Length 581;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  493;  Conservative   28;  Mismatches   43;  Indels   14;  Gaps    3;

Qy       4074 GGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAA 4133
Db        567 GGGTAAATATATATACAG--------TATGCTTACCTAGTCCTCCATGCCCGTGCAGCAA 516

Qy       4134 TATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCA 4193
Db        515 TATGGAGAGGAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCA 456

Qy       4194 TTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGG 4253
Db        455 TTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGG 396

Qy       4254 GTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMA 4313
Db        395 GTCACCAGAAAGGAAAGGAAAGGGAAATAGAAAGGAACCACCAAGCGGATATTGAAGCCA 336

Qy       4314 AAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGG 4373
Db        335 AAAGAGCCACAAGGCAGGACCCTCCATTAGAAATGCTTATAGAAGGAGCCCTAGTATAGG 276

Qy       4374 GTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCT 4433
Db        275 GTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGA---AGAAAYAGAATGGGGAACCT 219

Qy       4434 CACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTT 4490
Db        218 CACGAGGACATAGTTTCCTCCCCTCAGAATGGCCAGCCACCGAAGAAGGAAAAATACTTA 159

Qy       4491 TGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCA 4550
Db        158 TGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCATATCTTTCTCTTAGGCA 99

Qy       4551 TCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTA 4610
Db          98 TTAATAGCACCCATCAGATGGTCAAATTATTATTTACTAGACCAGGCCTTTTCAAAACTA 39

Qy       4611 TCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARA 4648
Db          38 TCAAGCAGATAGTCAAGGCCTGTGAAGTGTGCCAAAGA 1


RESULT 351
US-09-925-065A-787519/c
; Sequence 787519, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787519
;   LENGTH: 581
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-787519

  Query Match             6.0%;  Score 448.2;  DB 1;  Length 581;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  493;  Conservative   28;  Mismatches   43;  Indels   14;  Gaps    3;

Qy       4074 GGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAA 4133
Db        567 GGGTAAATATATATACAG--------TATGCTTACCTAGTCCTCCATGCCCGTGCAGCAA 516

Qy       4134 TATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCA 4193
Db        515 TATGGAGAGGAAGGGAATTCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCA 456

Qy       4194 TTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGG 4253
Db        455 TTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGG 396

Qy       4254 GTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMA 4313
Db        395 GTCACCAGAAAGGAAAGGAAAGGGAAATAGAAAGGAACCACCAAGCGGATATTGAAGCCA 336

Qy       4314 AAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGG 4373
Db        335 AAAGAGCCACAAGGCAGGACCCTCCATTAGAAATGCTTATAGAAGGAGCCCTAGTATAGG 276

Qy       4374 GTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCT 4433
Db        275 GTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGA---AGAAAYAGAATGGGGAACCT 219

Qy       4434 CACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTT 4490
Db        218 CACGAGGACATAGTTTCCTCCCCTCAGAATGGCCAGCCACCGAAGAAGGAAAAATACTTA 159

Qy       4491 TGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCA 4550
```

```
Db        158 TGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACCATATCTTTCTCTTAGGCA 99

Qy       4551 TCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTA 4610
              |  |||||||||||||:|||| ||||| |||||||||| ||||||||||||||||||||
Db         98 TTAATAGCACCCATCAGATGGTCAAATTATTATTTACTAGACCAGGCCTTTTCAAAACTA 39

Qy       4611 TCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARA 4648
              |||||||:|||:||| ||||||||||:|||||||:|:|
Db         38 TCAAGCAGATAGTCAAGGCCTGTGAAGTGTGCCAAAGA 1


RESULT 352
US-09-925-065A-948681/c
; Sequence 948681, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948681
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948681

  Query Match             6.0%;  Score 448;  DB 1;  Length 570;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  491;  Conservative   24;  Mismatches   52;  Indels   14;  Gaps    2;

Qy       4043 AYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
              |:|  ||||||| ||||||||||||||||  |  |           |||||||:|:|||
Db        570 ACTCAGGCACTAAAATTAGGAGAAGRAAAAGAGAGA----------AGACTCTAAGTAT 522

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
              |||:|||||||| ||||||||  ::|||||||:|:|||||||||||||||||||||:
Db        521 GCTTACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCT 462

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
              |||:|||||||||||:||| ||||||||||||||||:|||||||:||||:|||||||||
Db        461 GAGGGAACACCTATCAAACAACAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAA 402

Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
              ||||:|||||||||:|||||||||||||:|||||||||| |||||||||:||||||||||
Db        401 CCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATA 342

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
              |||:|||:  |||||| :||||| ||||:|||| |||  |||||||||| ||||||||
Db        341 AAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAGGCAGGACGCTCCATTA 282

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
              |||||||||| ||   ||| ||| || |||||||||| ||||||||||||||||||||
Db        281 GAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAG 222

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
              |||||||||| ||||| || |||||||| |||||||||    ||||| |||||||| |||
Db        221 TACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAGTTTCCTCCCCTCAGGA 162

Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
              || |||||||  ||||||| ||||||||  |||||||| ||| ||| || |||||||||
Db        161 AGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAACCAGTGAAAATTACTT 102

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
              |||||||||| ||||| |||||||||||||||  ||||||||||||:||||||||||
Db        101 AAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCA 42

Qy       4580 TTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARAT 4620
              |||||||| ||||||||||||||||||||||||||||:||
Db         41 TTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGAT 1


RESULT 353
US-09-925-065A-948681/c
; Sequence 948681, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948681
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948681

  Query Match             6.0%;  Score 448;  DB 1;  Length 570;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  491;  Conservative   24;  Mismatches   52;  Indels   14;  Gaps    2;

Qy       4043 AYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTAT 4102
              |:|  ||||||| ||||||||||||||||  |  |           |||||||:|:|||
Db        570 ACTCAGGCACTAAAATTAGGAGAAGRAAAAGAGAGA----------AGACTCTAAGTAT 522

Qy       4103 GCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCY 4162
              |||:|||||||| ||||||||||  ::|||||||:|:|||||||||||||||||||||:
Db        521 GCTTACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGGAATTCCTAACTTCT 462

Qy       4163 GAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAA 4222
              |||:|||||||||||:||| ||||||||||||||||:|||||||:||||:|||||||||
Db        461 GAGGGAACACCTATCAAACAACAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAA 402

Qy       4223 CCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATA 4282
              ||||:|||||||||:|||||||||||||:|||||||||| |||||||||:||||||||||
Db        401 CCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATA 342

Qy       4283 SAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTA 4342
              |||:|||:  |||||| :||||| ||||:|||| |||  |||||||||| ||||||||
Db        341 AAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAGGCAGGACGCTCCATTA 282

Qy       4343 GAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAG 4402
              |||||||||| ||   ||| ||| || |||||||||| ||||||||||||||||||||
Db        281 GAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAG 222

Qy       4403 TACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGA 4459
              |||||||||| ||||| || |||||||| |||||||||    ||||| |||||||| |||
Db        221 TACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAGTTTCCTCCCCTCAGGA 162
```

```
Qy       4460 CGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTT 4519
              || |||||||  |||||||  ||||||||  ||||||||  |||  |||  ||  |||||||||
Db        161 AGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAACCAGTGAAAATTACTT 102

Qy       4520 AAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCA 4579
              ||||||||||| |||||  |||||||||||||||  |||||||||||||:||||||||||||
Db        101 AAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATCA 42

Qy       4580 TTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARAT 4620
              ||||||||  ||||||||||||||||||||||||||||:||
Db         41 TTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGAT 1


RESULT 354
US-09-925-065A-945278/c
; Sequence 945278, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 945278
;   LENGTH: 577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-945278

  Query Match             6.0%;  Score 446.9;  DB 1;  Length 577;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  504;  Conservative   34;  Mismatches   37;  Indels   13;  Gaps    5;

Qy       3581 TGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTA 3640
              |||  ||||||||||::||| |||||||||||||||||||:||||||:||||||||||
Db        577 TGGCAGCAGTAGCAGTCTTAGGATCTGAAGCAGTTAAAATAAYACAGGGAAGAGATCTTA 518

Qy       3641 CTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGT 3700
              |||||||||||||||||  |:|| ||::|||||||||||||||||| ||||||||||||||
Db        517 CTGTGTGGACATCTCATAATGTAAACAGCATACTCACTGCTAAAAGAGACTTGTGGCTGT 458

Qy       3701 CAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGT 3760
              ||||||||: ||| |||||||:|||||||||||||    ||: ||||||||||||:|||||
Db        457 CAGACAACCATTTGCTTAAATATCAGGCTCTATTA----AAGAGCCAGTGCTGCAACTGT 402

Qy       3761 GCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAY 3820
              |||||||||||||||||| ||||:| |||||||||||||||| || |||||    ||:
Db        401 GCACTTGTGCAACTCTTAATCCAGCCACATTTCTTCCAGACAATAAAAAAAAG----AAC 346

Qy       3821 ATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTC 3880
              |||||||||||||  :| ||| |||| ||||| ||| ||  ||||||||||:||||:||||
Db        345 ATAACTGTCAACAGGTGATTGCTCACACCTACGCCGCTTAAGGGGACCTTCTAGAGGTTC 286

Qy       3881 CYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 3940
              |:|||||||||||:|||| |||||||||||||||||||||||||||||||||||||||||||
Db        285 CCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 227

Qy       3941 CTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
              ||||||||||: ||||| |||||| ||||||||||||||||:|| ||||||| ||||||
Db        226 CTTCGAAAAGCAGGGTACGCAGTGCTCAGTGATAATGGAATACTTCAAAGTAACCCCCTC 167

Qy       4001 ACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTA 4060
              |||||||||||||| |||:||||:|||||||||| || |:|||:|  ||||||||||||
Db        166 ACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAACAGACCTCATTCAGGCACTAGAATT- 108

Qy       4061 GGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
              |||||:|||||||||:|||||||||||||||| || |:|||||:|||||||| |||||
Db        107 --AGAAGGAAAAAGGGTAAATATATATACAGACACT-AATATGCTTACCTAGTCCTCCAT 51

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGA 4168
              | ||| |::|||||||:|:|||||||||||||||||: ||: |
Db         50 GACCACGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCAAGGAA 3


RESULT 355
US-09-925-065A-945278/c
; Sequence 945278, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 945278
;   LENGTH: 577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-945278

  Query Match             6.0%;  Score 446.9;  DB 1;  Length 577;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  504;  Conservative   34;  Mismatches   37;  Indels   13;  Gaps    5;

Qy       3581 TGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTA 3640
              |||  ||||||||||::||| |||||||||||||||||||:||||||:||||||||||
Db        577 TGGCAGCAGTAGCAGTCTTAGGATCTGAAGCAGTTAAAATAAYACAGGGAAGAGATCTTA 518

Qy       3641 CTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGT 3700
              |||||||||||||||||  |:|| ||::|||||||||||||||||| ||||||||||||||
Db        517 CTGTGTGGACATCTCATAATGTAAACAGCATACTCACTGCTAAAAGAGACTTGTGGCTGT 458

Qy       3701 CAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGT 3760
              ||||||||: ||| |||||||:|||||||||||||    ||: ||||||||||||:|||||
Db        457 CAGACAACCATTTGCTTAAATATCAGGCTCTATTA----AAGAGCCAGTGCTGCAACTGT 402

Qy       3761 GCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAY 3820
              |||||||||||||||||| ||||:| |||||||||||||||| || |||||    ||:
Db        401 GCACTTGTGCAACTCTTAATCCAGCCACATTTCTTCCAGACAATAAAAAAAAG----AAC 346

Qy       3821 ATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTC 3880
              |||||||||||||  :| ||| |||| ||||| ||| ||  ||||||||||:||||:||||
Db        345 ATAACTGTCAACAGGTGATTGCTCACACCTACGCCGCTTAAGGGGACCTTCTAGAGGTTC 286

Qy       3881 CYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 3940
              |:|||||||||||:|||| |||||||||||||||||||||||||||||||||||||||||||
Db        285 CCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGA 227

Qy       3941 CTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTC 4000
              ||||||||||: ||||| |||||| ||||||||||||||||:|| ||||||| ||||||
```

```
Db        226 CTTCGAAAAGCAGGGTACGCAGTGCTCAGTGATAATGGAATACTTCAAAGTAACCCCCTC 167

Qy       4001 ACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTA 4060
             ||||||||||||||| |||:||||:|||||||||| || |:|||:|  ||||||||||||
Db        166 ACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAACAGACCTCATTCAGGCACTAGAATT- 108

Qy       4061 GGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCAT 4120
              |||||:||||||||:|||||||||||||||| || |:||||||:|||||||| |||||
Db        107 --AGAAGGAAAAAGGGTAAATATATATACAGACACT-AATATGCTTACCTAGTCCTCCAT 51

Qy       4121 GCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGA 4168
              | ||| |::||||||||:|:|||||||||||||||||||||||: ||: |
Db         50 GACCACGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCAAGGAA 3


RESULT 356
US-09-925-065A-789246/c
; Sequence 789246, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 789246
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-789246

  Query Match             6.0%;  Score 446.8;  DB 1;  Length 560;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  490;  Conservative   41;  Mismatches   29;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
             ||||||||||||||||||||||||||||||||||||          |||:  ||||||||||
Db        560 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 510

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
             |||||||:|||||||||||||||||| ||||||:|||| ||||||||||||||||||||||
Db        509 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 450

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
             |||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| |||||||||
Db        449 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 390

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
             | || |:||: ||::|||||||||||||| |||||||||||||||:|||||:|||||
Db        389 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 330

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
             ||||||:|||||||||          ||||||| |:|:|||||:||||||||||||||||
Db        329 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 279

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
             |||||||||||:||||:|||||||||||| |||||:|||||||| |||||||::|||||
Db        278 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 219

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
              || ||||||||||||||||||:|||||||||||||||||||||:|::|||||||||:|
Db        218 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 159

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
             ||:|| :|||||||||:||||||| ||||| |:  | :|||||||||:|||||: ||
Db        158 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 100

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
             || |||||||||||||||||||||||||||||||| ||:|||||||||:|:|||||  :::
Db         99 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 40

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAA 3404
              |||||||||||||| |||||||||||||| ||| || ||
Db         39 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAA 1


RESULT 357
US-09-925-065A-789246/c
; Sequence 789246, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 789246
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-789246

  Query Match             6.0%;  Score 446.8;  DB 1;  Length 560;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  490;  Conservative   41;  Mismatches   29;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
             ||||||||||||||||||||||||||||||||||||          |||:  ||||||||||
Db        560 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 510

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
             |||||||:|||||||||||||||||| ||||||:|||| ||||||||||||||||||||||
Db        509 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 450

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
             |||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| |||||||||
Db        449 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 390

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
             | || |:||: ||::|||||||||||||| |||||||||||||||:|||||:|||||
Db        389 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 330

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
             ||||||:|||||||||          ||||||| |:|:|||||:||||||||||||||||
Db        329 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 279

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
             |||||||||||:||||:|||||||||||| |||||:|||||||| |||||||::|||||
Db        278 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 219
```

```
Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
Db        218 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 159

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
Db        158 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 100

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
Db          99 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 40

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAA 3404
Db          39 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAA 1


RESULT 358
US-09-925-065A-789955/c
; Sequence 789955, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 789955
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-789955

  Query Match             6.0%;  Score 446.8;  DB 1;  Length 562;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  490;  Conservative   41;  Mismatches   29;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
Db        562 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 512

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
Db        511 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 452

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
Db        451 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 392

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
Db        391 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 332

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
Db        331 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 281

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
Db        280 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 221
```

```
Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
Db        220 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 161

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
Db        160 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 102

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
Db        101 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 42

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAA 3404
Db         41 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAA 3


RESULT 359
US-09-925-065A-789955/c
; Sequence 789955, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 789955
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-789955

  Query Match             6.0%;  Score 446.8;  DB 1;  Length 562;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  490;  Conservative   41;  Mismatches   29;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
Db        562 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 512

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
Db        511 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 452

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
Db        451 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 392

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
Db        391 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 332

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
Db        331 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 281

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
```

```
Db        280 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 221

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
             || ||||||||||||||||||||:||||||||||||||||||||:|::||||||||||:|
Db        220 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 161

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
             ||:|| :|||||||||:||||||| ||||| |:  | :||||||||||:|||||: ||
Db        160 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 102

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
             || |||||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::
Db        101 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 42

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAA 3404
             |||||||||||||| ||||||||||||| ||| || ||
Db         41 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCCAAA 3

RESULT 360
US-09-925-065A-756176/c
; Sequence 756176, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 756176
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-756176

  Query Match             5.9%;  Score 445.4;  DB 1;  Length 558;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  488;  Conservative   41;  Mismatches   28;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
             ||||||||||||||||||||||||||||||||||||      |||:  ||||||||||
Db         558 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 508

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
             |||||||:|||||||||||||||||| ||||||:|||| ||||||||||||||||||||||
Db         507 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 448

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
             ||||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| ||||||||||
Db         447 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 388

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
             | || |:||: ||::|||||||||||||||| ||||||||||||||||||:|||||:|||||
Db         387 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 328

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
             ||||||:|||||||||          ||||||| |:|:|||||:|||||||||||||||||
Db         327 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 277

Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
```

```
             ||||||||||||:||||:|||||||||| |||||:|||||||| |||||||::|||||
Db         276 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 217

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
             || ||||||||||||||||||||:||||||||||||||||||||:|::||||||||||:|
Db         216 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 157

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
             ||:|| :|||||||||:||||||| ||||| |:  | :||||||||||:|||||: ||
Db         156 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 98

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
             || |||||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::
Db          97 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 38

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
             |||||||||||||| ||||||||||||-||| ||
Db          37 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCC 2

RESULT 361
US-09-925-065A-756176/c
; Sequence 756176, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 756176
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-756176

  Query Match             5.9%;  Score 445.4;  DB 1;  Length 558;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  488;  Conservative   41;  Mismatches   28;  Indels   20;  Gaps    4;

Qy       2825 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884
             ||||||||||||||||||||||||||||||||||||      |||:  ||||||||||
Db         558 TCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT---------CCAATGTCTCAACTCAC 508

Qy       2885 CTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGC 2944
             |||||||:|||||||||||||||||| ||||||:|||| ||||||||||||||||||||||
Db         507 CTGGACTGTTTTACCCCAAGGGTTCAAGGATAGCCCCCGTCTATTTGGCCAGGCATTAGC 448

Qy       2945 CCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTT 3004
             ||||||||||||:||:|: ||||| ||||||| |||||| ||||:|||:| ||||||||||
Db         447 CCAAGACTTGAGCCAATTGTCATAGCTGGACATTCTTGTTCTTCAGTATGGGGATGATTT 388

Qy       3005 ACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTT 3064
             | || |:||: ||::|||||||||||||||| ||||||||||||||||||:|||||:|||||
Db         387 AATTATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTT 328

Qy       3065 CCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGT 3124
             ||||||:|||||||||          ||||||| |:|:|||||:|||||||||||||||||
Db         327 CCTCGCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGT 277
```

```
Qy       3125 TAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA 3184
              ||||||||||||:||||:|||||||||||| |||||:||||||||| |||||||::|||||
Db        276 TAAATACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAAYGTATCCA 217

Qy       3185 GCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRT 3244
              || ||||||||||||||||||||:|||||||||||||||||||||:|::|||||||||:|
Db        216 GCGTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCAT 157

Qy       3245 AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATA 3304
              ||:|| :||||||||||:||||||| ||||| |:  | :||||||||||:|||||: ||
Db        156 AACAGCCTTCTGCCGAATATGGATT-CCCAGATACAGTGAAATAGCCAGGCCATTATGTA 98

Qy       3305 CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364
              || |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::
Db          97 CATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAG 38

Qy       3365 AAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
              ||||||||||||||| ||||||||||||||| ||| ||
Db          37 AAGTGGCTTTCCAGG-CCCTAAAGAAGGCCCTAACCC 2
```

RESULT 362
US-09-864-761-7501/c
; Sequence 7501, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 7501
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139245.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
US-09-864-761-7501

Query Match 5.9%; Score 445.1; DB 1; Length 559;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 492; Conservative 13; Mismatches 42; Indels 19; Gaps 2;

```
Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
              ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
              |:||||||| ||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
              |||||||||||||  |||||||||||||||||||||||:|| |||           |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              |||||||||||| ||| |||||||| ||||| |||||||| || |||||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| ||||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              |         ||    |||||||||||||||  ||| |||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 363
US-10-027-632-192839
; Sequence 192839, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

```
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 192839
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-192839

  Query Match             5.9%;  Score 439.6;  DB 1;  Length 598;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  530;  Conservative    2;  Mismatches   61;  Indels   10;  Gaps    5;

Qy       1583 GACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGAC 1642
             ||||||  |||||||||||||||||||||    |||||||||||||||| ||||||||||
Db             1 GACTCTATTGAAGGCCAACTAATCTTAAAAGATAAGTTTATCACTCAGCCAGCTGCAGAC 60

Qy       1643 ATTAG--AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATT 1700
             || ||  ||||||||||||||||||||  |||||||||||||||||||||||||||||| |
Db            61 ATCAGAAAAAAAAACTTCAAAAGTCTGCTTTAGGCCCGGAGCAAAACTTAGAAACCCTACT 120

Qy       1701 GAACTTGGCAACYTCGGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAA 1760
             ||||||||||||:||||||||||||||||||||||||||||||||||||| ||||||||||
Db            121 GAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAACAGGACAG 180

Qy       1761 ACGGGATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAG 1820
             | | || ||| |||||||||||||||||||||||||| |||||    ||||  ||||||||||
Db            181 ATGAGA-TAAGAAAAAGGCCACCGCTTTAGTCATGGCCCTC---CAAGCTGACTTTGGAG 236

Qy       1821 GCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCG 1880
             |||||||||    |||||  |||||||||| ||||||||||||||||||||||||||||||
Db            237 GCTCTGGAACGTGGAAAGCCTGGGCAAATCGAATGCCTAATAGGGCTTGCTTCCAGTGCA 296

Qy       1881 GTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCAT 1940
             |||||||||||||||||||| |||||||||| || ||||| | ||| ||||| | ||||||
Db            297 GTCTACAAGGACACTTTAAAGAAGATTGTCGAA-TAGAAATGAGCTGCCCCCTTGTCCAT 355

Qy       1941 GCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTG 2000
             |||||||||  |||||||||||| |||||||||||||||||||| || |  |||||| ||||
Db            356 GCCCCTTATCTCAAGGGAATCATTGGAAGGCCCACTGCCCCAAGGAATGAAGGTCGTTTG 415

Qy       2001 AGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGC 2060
             | |||||||||||||||| |||||||||||||||||||||||||||||| |||||||| ||
Db            416 AATCAGAAGCCACTAACCTGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGC 475

Qy       2061 CATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG 2120
             |  ||||| ||||||||||||||||||:|  |||||||||||||||||||||||||||| |
Db            476 CTGCCCATACCATCACCCTCACAGAGCYCCAGGTATGCTTGACCATTGAGGGCCAGGAGG 535

Qy       2121 GT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAAC 2177
               |   |||||||||||||||||   || ||||| | ||||||| | ||||||   |||||
Db            536 TTAACTGTCTCCTGGACACTGGCATGGCCTTCTCAATCTTACTTTCCTGTCCTATACAAC 595

Qy       2178 TGT 2180
             |||
Db            596 TGT 598


RESULT 364
US-10-027-632-192839
; Sequence 192839, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 192839
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-192839

  Query Match             5.9%;  Score 439.6;  DB 1;  Length 598;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  530;  Conservative    2;  Mismatches   61;  Indels   10;  Gaps    5;

Qy       1583 GACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGAC 1642
             ||||||  |||||||||||||||||||||    |||||||||||||||| ||||||||||
Db             1 GACTCTATTGAAGGCCAACTAATCTTAAAAGATAAGTTTATCACTCAGCCAGCTGCAGAC 60

Qy       1643 ATTAG--AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATT 1700
             || ||  ||||||||||||||||||||  |||||||||||||||||||||||||||||| |
Db            61 ATCAGAAAAAAAAACTTCAAAAGTCTGCTTTAGGCCCGGAGCAAAACTTAGAAACCCTACT 120

Qy       1701 GAACTTGGCAACYTCGGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAA 1760
             ||||||||||||:||||||||||||||||||||||||||||||||||||| ||||||||||
Db            121 GAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAACAGGACAG 180

Qy       1761 ACGGGATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAG 1820
             | | || ||| |||||||||||||||||||||||||| |||||    ||||  ||||||||||
Db            181 ATGAGA-TAAGAAAAAGGCCACCGCTTTAGTCATGGCCCTC---CAAGCTGACTTTGGAG 236

Qy       1821 GCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCG 1880
             |||||||||    |||||  |||||||||| ||||||||||||||||||||||||||||||
Db            237 GCTCTGGAACGTGGAAAGCCTGGGCAAATCGAATGCCTAATAGGGCTTGCTTCCAGTGCA 296

Qy       1881 GTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCAT 1940
             |||||||||||||||||||| |||||||||| || ||||| | ||| ||||| | ||||||
Db            297 GTCTACAAGGACACTTTAAAGAAGATTGTCGAA-TAGAAATGAGCTGCCCCCTTGTCCAT 355

Qy       1941 GCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTG 2000
             |||||||||  |||||||||||| |||||||||||||||||||| || |  |||||| ||||
Db            356 GCCCCTTATCTCAAGGGAATCATTGGAAGGCCCACTGCCCCAAGGAATGAAGGTCGTTTG 415

Qy       2001 AGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGC 2060
             | |||||||||||||||| |||||||||||||||||||||||||||||| |||||||| ||
Db            416 AATCAGAAGCCACTAACCTGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGC 475

Qy       2061 CATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG 2120
             |  ||||| ||||||||||||||||||:|  |||||||||||||||||||||||||||| |
Db            476 CTGCCCATACCATCACCCTCACAGAGCYCCAGGTATGCTTGACCATTGAGGGCCAGGAGG 535

Qy       2121 GT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAAC 2177
               |   |||||||||||||||||   || ||||| | ||||||| | ||||||   |||||
Db            536 TTAACTGTCTCCTGGACACTGGCATGGCCTTCTCAATCTTACTTTCCTGTCCTATACAAC 595
```

```
Qy       2178 TGT 2180
              |||
Db        596 TGT 598


RESULT 365
US-09-925-065A-919226
; Sequence 919226, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919226
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919226

  Query Match             5.8%;  Score 437.8;  DB 1;  Length 548;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  472;  Conservative   53;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3020 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGG 3079
              ||||||||||||||||||||||||||||||||||||:|||||:||||||| ||:|||||||
Db          1 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTTGCCACCTGTGG 60

Qy       3080 CTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT 3139
              |||||:|||||||||||:| |:|||||:||||||||| | |||| ||||||||||||:||
Db         61 CTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCT 119

Qy       3140 AARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTA 3199
              ||:||||||||||||||||:|||||||||||||||||:::|||||||| ||||||||||
Db         120 AAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTA 179

Qy       3200 TCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCG 3259
              ||||||||:|||||||||||||||||||:|::||| |||||:|| :|||:||||||||
Db         180 TCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCG 239

Qy       3260 AAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAA 3319
              ||:|||||||| |||||||: |||:||||||| |||:|||||:|||||:||||| ||||||
Db         240 AATATGGATT-CCCAGGTAMGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAA 298

Qy       3320 CTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGG 3379
              |||||||||||||||||||||||:||||||||||:|:||||||:::||| ||||||||||
Db         299 CTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGG 358

Qy       3380 CCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGA 3436
              |||  ||  | | |||  ||||||||:||||||   ||||:||||||||:||||||||
Db         359 CCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGA 416

Qy       3437 CTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCC 3496
              |||| :||:|||::|||||| |||||||||:||:||||||:||||||||||||||||:|||
Db         417 CTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCC 475

Qy       3497 RAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGG 3556
              :|||||: ||| |||||||::||||::|||||||:||||||||:|||||||||||||||||
Db         476 AAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGG 535

Qy       3557 GTTGRCYTCATTG 3569
              ||||:|:||||||
Db         536 GTTGGCCTCATTG 548


RESULT 366
US-09-925-065A-919226
; Sequence 919226, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919226
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919226

  Query Match             5.8%;  Score 437.8;  DB 1;  Length 548;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  472;  Conservative   53;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3020 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGG 3079
              ||||||||||||||||||||||||||||||||||||:|||||:||||||| ||:|||||||
Db          1 TTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTTGCCACCTGTGG 60

Qy       3080 CTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT 3139
              |||||:|||||||||||:| |:|||||:||||||||| | |||| ||||||||||||:||
Db         61 CTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCT 119

Qy       3140 AARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTA 3199
              ||:||||||||||||||||:|||||||||||||||||:::|||||||| ||||||||||
Db         120 AAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTA 179

Qy       3200 TCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCG 3259
              ||||||||:|||||||||||||||||||:|::||| |||||:|| :|||:||||||||
Db         180 TCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCG 239

Qy       3260 AAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAA 3319
              ||:|||||||| |||||||: |||:||||||| |||:|||||:|||||:||||| ||||||
Db         240 AATATGGATT-CCCAGGTAMGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAA 298

Qy       3320 CTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGG 3379
              |||||||||||||||||||||||:||||||||||:|:||||||:::||| ||||||||||
Db         299 CTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGG 358

Qy       3380 CCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGA 3436
              |||  ||  | | |||  ||||||||:||||||   ||||:||||||||:||||||||
Db         359 CCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGA 416

Qy       3437 CTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCC 3496
              |||| :||:|||::|||||| |||||||||:||:||||||:||||||||||||||||:|||
Db         417 CTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCC 475

Qy       3497 RAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGG 3556
```

```
              :|||||: ||| |||||||::||||::||||||:||||||||:||||||||||||||||||
Db        476 AAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGG 535

Qy       3557 GTTGRCYTCATTG 3569
              ||||:|:||||||
Db        536 GTTGGCCTCATTG 548


RESULT 367
US-09-864-761-32803
; Sequence 32803, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32803
;   LENGTH: 491
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL122001.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 4.00e-26
;   OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
US-09-864-761-32803

  Query Match             5.8%;  Score 433.2;  DB 1;  Length 491;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  423;  Conservative   38;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
              ||:|||||||:|||||||||||||:||:|||::|||||||||||||||:||:| ||||
Db          1 AGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAGAAAAAAACAGGAATAACTCT 60

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
              :|||||||||||||||:|||:||| |:||||||||||||::||||: |||||:||||||
Db         61 AGGAGTCCTTACACAGGTCCAAGGAACGAGCTTGCAACCCGTGGCAGACCTGAGTAAGGA 120

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
              ||:|||||||||||||||||||||:|:||||||||||:||||| |||||||||||||| |
Db         121 AATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTGGCAGTAGCAAT 180

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
              ::|||||||||||||||||||||||||||||||| :|||||||||||||||||||||||||
Db         181 CTTAGTATCTGAAGCAGTTAAAATAATACAGGAAAGAGATCTTACTGTGTGGACATCTCA 240

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
              |||:|||||::|||||||||||||||||||||| |||||||||||||| |: ||||||
Db         241 TGATGTGAACAGCATACTCACTGCTAAAGGAGACCTGTGGCTGTCAGACAGCCATTTACT 300

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
              |||||: |||| ||||||||||||:||||||||||:|||||||||||| ||||||||
Db         301 TAAATACCAGGGTCTATTACTTGAAGGGCCAGTGCTGCAACTGTGCACTTGGGCAACTCT 360

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
              ||||||||:| |||||||||||||||| |||| ||||||||:||: ||||||||||| :|
Db         361 TAACCCAGCCACATTTCTTCCAGACATTGAAAAAAAGATAGAACGTAACTGTCAACAGGT 420

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
              ||| |||||||||||||| ||||||| |||||:||||:|||||:||||||||||:||
Db         421 GATTGCTCAAACCTATGCCGCTCGAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGA 480

Qy       3897 CCT 3899
              |||
Db         481 CCT 483


RESULT 368
US-10-301-480-473577/c
; Sequence 473577, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 473577
;   LENGTH: 551
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-473577

  Query Match             5.7%;  Score 430.3;  DB 1;  Length 551;
  Best Local Similarity   82.1%;  Pred. No. 0;
  Matches  453;  Conservative   47;  Mismatches   49;  Indels    3;  Gaps    2;

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
              ||||||||||:|| :|||:|||||| |||:|||||| |||||||:  ||:|||||||||
Db         551 TTCCTTGGCATAGCAGGTTTCTGCTGAATATGGATT-CCCAGGTACCGCAAAATAGCCAC 493

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
              :|| ||:|||||:||||||| | |||||||| ||||||||||||||||:|||| ||||:
```

```
Db        492 ACCACTATATACACTAATTAAAGGAACTCAGACAGCCAATACCCATTTAGTAAGCTGGAC 433

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
              | ||||||:::||| ||||||||||| |  ||  | | |||  |||  |||:|||||||
Db        432 ATCTGAAGCAGAAGAGGCTTTCCAGGCACTAAAGAAGGCCCT--AACTTAAGCCCCAGTG 375

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
              |||||:|| |||||:||:|||||||||||:||:|||::|||||||||||||| :||:|||
Db        374 TTAAGCTTCCCAACAGGRCAAGACTTTTCTTTATATGTCACAGAAAAAAACAAGAATAGC 315

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
              || :||||||||||| ||:|| :|| ||:|||||| ||||::||||::||||||:|||
Db        314 TCCAGGAGTCCTTACATAGGTCTGAGAGATGAGCTTGGAACCCATGGCATACCTGAGTAA 255

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
              |||||:||||| ||||||||||||||| |:||||||||||:|| ||||||  |||||| |
Db        254 GGAAATTGATGCAGTGGCAAAGGGTTGTCCTCATTGTTTATGGATAGTGGCAGCAGTAAC 195

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATC 3653
              | |::  |||||| |||||||||||||||| ||:|| ||||||||||||||||||||
Db        194 AATCTAGGTATCTTTAGCAGTTAAAATAATACAAGGAAGGGATCTTACTGTGTGGACATC 135

Qy       3654 TCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTT 3713
              ||||||:|||||::|||||||||||||||||||||||||||||||||||||||||: |||
Db        134 TCATGATGTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTTTTT 75

Qy       3714 ACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAAC 3773
              ||||||||:|||||||||||||| |||||: || ||||||| :| || |||||||||||||
Db         74 ACTTAAATATCAGGCTCTATTATTTGAAGTGCTAGTGCTGTGAATGCGCACTTGTGCAAC 15

Qy       3774 TCTTAACCCAGY 3785
              ||||||||||:
Db         14 TCTTAACCCAGC 3


RESULT 369
US-10-301-480-1086986/c
; Sequence 1086986, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1086986
;   LENGTH: 551
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1086986

  Query Match             5.7%;  Score 430.3;  DB 1;  Length 551;
  Best Local Similarity   82.1%;  Pred. No. 0;
  Matches  453;  Conservative   47;  Mismatches   49;  Indels    3;  Gaps    2;

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
              |||||||||:|| :|||:|||||| |||:|||||| |||||||:  ||:|||||||||
Db         551 TTCCTTGGCATAGCAGGTTTCTGCTGAATATGGATT-CCCAGGTACCGCAAAATAGCCAC 493

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
               :|| ||:|||||:|||||| | |||||||| ||||||||||||||:|||| |||||:
Db         492 ACCACTATATACACTAATTAAAGGAACTCAGACAGCCAATACCCATTTAGTAAGCTGGAC 433

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
              | ||||||:::||| ||||||||||| |  ||  | | |||  |||  |||:|||||||
Db         432 ATCTGAAGCAGAAGAGGCTTTCCAGGCACTAAAGAAGGCCCT--AACTTAAGCCCCAGTG 375

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
              |||||:|| |||||:||:|||||||||||:||:|||::|||||||||||||| :||:|||
Db         374 TTAAGCTTCCCAACAGGRCAAGACTTTTCTTTATATGTCACAGAAAAAAACAAGAATAGC 315

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
              || :||||||||||| ||:|| :|| ||:|||||| ||||::||||::||||||:|||
Db         314 TCCAGGAGTCCTTACATAGGTCTGAGAGATGAGCTTGGAACCCATGGCATACCTGAGTAA 255

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGC 3593
              |||||:||||| ||||||||||||||| |:||||||||||:|| ||||||  |||||| |
Db         254 GGAAATTGATGCAGTGGCAAAGGGTTGTCCTCATTGTTTATGGATAGTGGCAGCAGTAAC 195

Qy       3594 AGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATC 3653
              | |::  |||||| |||||||||||||||| ||:|| ||||||||||||||||||||
Db         194 AATCTAGGTATCTTTAGCAGTTAAAATAATACAAGGAAGGGATCTTACTGTGTGGACATC 135

Qy       3654 TCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTT 3713
              ||||||:|||||::|||||||||||||||||||||||||||||||||||||||||: |||
Db         134 TCATGATGTGAACAGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTTTTT 75

Qy       3714 ACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAAC 3773
              ||||||||:|||||||||||||| |||||: || ||||||| :| || |||||||||||||
Db          74 ACTTAAATATCAGGCTCTATTATTTGAAGTGCTAGTGCTGTGAATGCGCACTTGTGCAAC 15

Qy       3774 TCTTAACCCAGY 3785
              ||||||||||:
Db          14 TCTTAACCCAGC 3


RESULT 370
US-10-029-386-20359
; Sequence 20359, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20359
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022067.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.44
;   OTHER INFORMATION: SWISSPROT HIT: P10272, EVALUE 1.00e-38
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF989696.1, EVALUE 1.00e-92
US-10-029-386-20359

  Query Match             5.7%;  Score 429.5;  DB 1;  Length 566;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  507;  Conservative   11;  Mismatches   46;  Indels   13;  Gaps    5;

Qy       2411 GAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGC 2470
              |||||||||||||| |||| |||||||||||||||||||||||||||| ||||||||||
Db            1 GAAGGAATTAATCTTGAAGACTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGC 59

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
              |||||||||||||||||||||||||||||  | ||||||||| ||||||||||||||||||
Db           60 CCGTCCTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCCACCAAAGGCAGTACCC 119
```

```
Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGG 2590
             ||| |||||| |||| |||||||||| |||||||||    |||||||| ||||||||||||||
Db        120 CCTTAGACCCGAGGCTCAACAAGGACTCCAAAAGA---TTAAGGACCTAAAAGCCCAAGG 176

Qy       2591 CTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGA 2650
             | |||||||| |||||| |  |||| || |||| |         | |   |  | |||||
Db        177 CCTAGTAAAAGCATGCAATAGCCCCTACAATAATCC-------AACTTTAGGAGTACAGA 229

Qy       2651 AACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGT 2710
             |||||||||||||||| |||||||||||||||||||||||||| |||| |   |||
Db        230 AACCCAGTGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGTCACTGT 287

Qy       2711 CCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGC 2770
             || | |||||| |||||||||| |||||||| | ||::||||||||||||||||||||||
Db        288 CCCTCTATACCTAGCTGTACCTAACCCTTATATTCTGCTTTCCCAAATACCAGAGGAAGC 347

Qy       2771 AGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGA 2830
             |||||||||||||: |||||||||:|||||||||| ||||||||||||||||||||||||
Db        348 AGAGTGGTTTACAGACCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGA 407

Qy       2831 CTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGAC 2890
             |||||||||||| |||||||||||||| ||||||||||:  |||||||||||||||||||
Db        408 CTCTCAATTCTTATTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAACTCACCTGGAC 467

Qy       2891 TRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGA 2950
             |:||| ||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        468 TGTTTCACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGA 527

Qy       2951 CTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTT 2987
             ||||||:| :|::|||||||||| ||||||||||||
Db        528 CTTGAGCCGGTTCTCATACCTGGGCACTCTTGTCCTT 564
```

RESULT 371
US-09-925-065A-296722/c
; Sequence 296722, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 296722
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-296722

Query Match 5.7%; Score 429.3; DB 1; Length 523;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 445; Conservative 53; Mismatches 25; Indels 5; Gaps 3;

```
Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:||||||||| |:|||||||||||||:|||||||| ||:|:|:|||||:||||||
Db        523 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 464

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||||:||||:||||||||||||||||:|||| ||||| ||
Db        463 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 404

Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
             ||: :||| ||||||||||||||||||||||:|||||| |||||||||||||:|::
Db        403 AAACCTATCTAGCCTATACTGGCTTATCCTCATCCCAAAACACTAAAGCAACTAAGAGGG 344

Qy       3234 TTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAG 3293
             |||||||||:||| |||:||||||||||:||||||| |||||||:  ||:|||||||||
Db        343 TTCCTTGGCATAAAAGGCTTCTGCCGAATATGGATT-CCCAGGTACAGCGAAATAGCCAG 285

Qy       3294 GYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAY 3353
             |:|||||:|||||:||||||||:|||||||||||||||||||||||||:|||||||||:
Db        284 GCCATTATATACACTAATTAAGRAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGAT 225

Qy       3354 AMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTG 3413
             |:||||||:::|||||||||||||||||||  ||  | | |||  ||||||| :|||||||
Db        224 ACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTG 167

Qy       3414 TTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGC 3473
             |||||:|||||||:| || ||||||||:||:|||::||||||  ||||||||:||:|||
Db        166 TTAAGCTTGCCAACAGTGCTAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGC 109

Qy       3474 TCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAA 3533
             |||:|||||||||||||:|| :|||||: |||||||||::||||::|| |||:|||
Db        108 TCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCTTGCAACCCGTGGCATACTTGAGTAA 49

Qy       3534 GGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
             |||||:|||||||||| ||||||||||:|:|||||||||: ||||||
Db         48 GGAAACTGATGTAGTGACAAAGGGTTGGCCTCATTGTTTACTGGTAGT 1
```

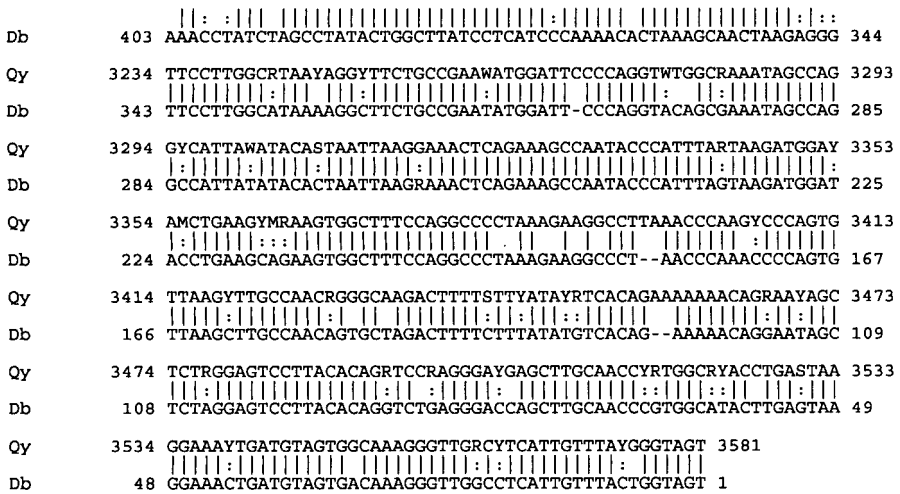
RESULT 372
US-09-925-065A-296722/c
; Sequence 296722, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 296722
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-296722

Query Match 5.7%; Score 429.3; DB 1; Length 523;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 445; Conservative 53; Mismatches 25; Indels 5; Gaps 3;

```
Qy       3054 CTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGC 3113
             |||||:||||||||| |:|||||||||||||:|||||||| ||:|:|:|||||:||||||
Db        523 CTCTTAAATTTCCTCACCACCTGTGGCTACAAGGTTTCCATACCAAAGGCTCAGCTCTGC 464

Qy       3114 TCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAG 3173
             |||||||||||||||||||||||:||||:||||||||||||||||:|||| ||||| ||
Db        463 TCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCTTCAGTAAG 404

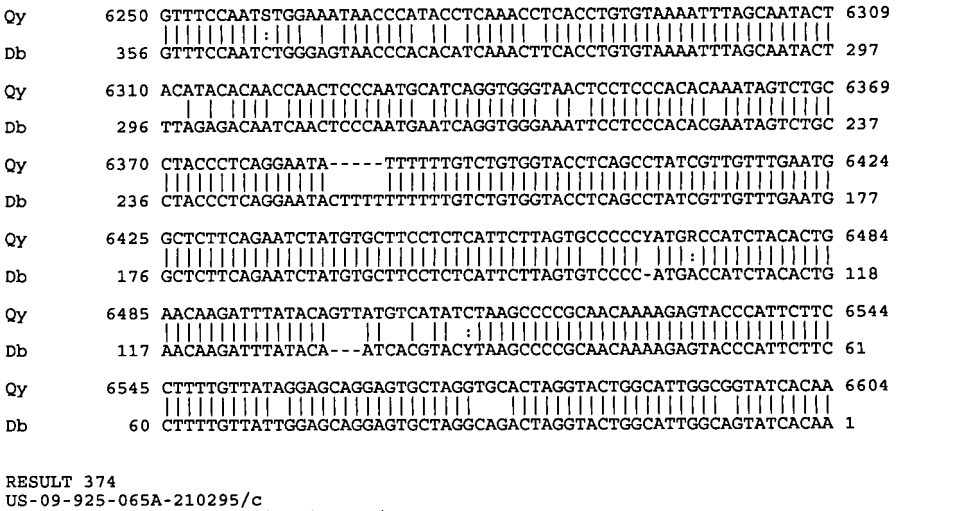
Qy       3174 GAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRR 3233
```



RESULT 373
US-09-925-065A-210295/c
; Sequence 210295, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210295
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-210295

Query Match 5.7%; Score 427.7; DB 1; Length 532;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 491; Conservative 4; Mismatches 32; Indels 13; Gaps 4;





RESULT 374
US-09-925-065A-210295/c
; Sequence 210295, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210295
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-210295

Query Match 5.7%; Score 427.7; DB 1; Length 532;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 491; Conservative 4; Mismatches 32; Indels 13; Gaps 4;



```
Qy       6310 ACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGC 6369
Db        296 TTAGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGAATAGTCTGC 237

Qy       6370 CTACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATG 6424
Db        236 CTACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATG 177

Qy       6425 GCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTG 6484
Db        176 GCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTG 118

Qy       6485 AACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTC 6544
Db        117 AACAAGATTTATACA---ATCACGTACYTAAGCCCCGCAACAAAAGAGTACCCATTCTTC 61

Qy       6545 CTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAA 6604
Db          60 CTTTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAA 1
```

RESULT 375
US-09-925-065A-857674
; Sequence 857674, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 857674
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-857674

Query Match 5.7%; Score 426.9; DB 1; Length 558;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 493; Conservative 13; Mismatches 49; Indels 9; Gaps 4;

```
Qy       6779 TAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG---AGAAAGTTR 6835
Db          1 TAGGGGAAAGATGCTGTTATTATGTTAATCAATCCAGAATCGTCACTGAGAAGAAAGYTA 60

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGG 6895
Db         61 AAAAAATTCGAGATCGCATACAACGTAGAGCAGGGGAGCTTCAAAACACCGGACCCTGGG 120

Qy       6896 GCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAA 6955
Db        121 GCCCCCTCAGCCAATGGATGCCCTGGATTCTCCCCTCCTTAGGACCTCTAGCAGCTATAA 180

Qy       6956 TATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCA 7015
Db        181 TATTGTTACTCCTCTTTGGACCCTGTATC-TTAACCTCCTTGTTAAGTTTGTCTCTTCCA 239

Qy       7016 GAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACC 7075
Db        240 GAATCAAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAG----TCC 295

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
Db        296 ATGGACCCCTGGACCGGCCTGCTAGCCTATGCTCTGATGTTAATGACATCGAAGGCAT-C 354

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
Db        355 CTCCTGAGAAAATCTCAACTGCACAACCGCTACTACACCCCAGTTCAGCAGGAAGCAGTT 414

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
Db        415 AGAGTGGTCGCTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 474

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
Db        475 TGAGAGACAGGACTAGCTGGATTTCCTAAGCCAACTAAGAATCCCTAAGCCTACCTGGGA 534

Qy       7316 AGGTGACCACATCCACCTTTAAAC 7339
Db        535 AGGTGACCACATCCACCTTTAAAC 558
```

RESULT 376
US-09-925-065A-857674
; Sequence 857674, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 857674
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-857674

Query Match 5.7%; Score 426.9; DB 1; Length 558;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 493; Conservative 13; Mismatches 49; Indels 9; Gaps 4;

```
Qy       6779 TAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG---AGAAAGTTR 6835
Db          1 TAGGGGAAAGATGCTGTTATTATGTTAATCAATCCAGAATCGTCACTGAGAAGAAAGYTA 60

Qy       6836 AAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGG 6895
Db         61 AAAAAATTCGAGATCGCATACAACGTAGAGCAGGGGAGCTTCAAAACACCGGACCCTGGG 120

Qy       6896 GCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAA 6955
Db        121 GCCCCCTCAGCCAATGGATGCCCTGGATTCTCCCCTCCTTAGGACCTCTAGCAGCTATAA 180

Qy       6956 TATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCA 7015
Db        181 TATTGTTACTCCTCTTTGGACCCTGTATC-TTAACCTCCTTGTTAAGTTTGTCTCTTCCA 239

Qy       7016 GAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACC 7075
```

```
              ||||| |||||||:|| ||||||||||  |   | ||| || || |||   ||     ||
Db        240 GAATCAAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAG----TCC 295

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              |||||||||||||||||:||||| | | |||||||||||||||||| ||||||  |
Db        296 ATGGACCCCTGGACCGGCCTGCTAGCCTATGCTCTGATGTTAATGACATCGAAGGCAT-C 354

Qy       7136 CTCCTGAGGAAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||||||| |||||||| |||||||||| ||||||| ||||| ||||||||||||||||||
Db        355 CTCCTGAGAAAATCTCAACTGCACAACCGCTACTACACCCCAGTTCAGCAGGAAGCAGTT 414

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              |||| |||:|  |||||||||||||||||||||| ||||||||||||||| |||||||
Db        415 AGAGTGGTCGCTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 474

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              |||||||||||||||||||||||||||| ||  |:|||||||||:|||||||| :|||||
Db        475 TGAGAGACAGGACTAGCTGGATTTCCTAAGCCAACTAAGAATCCCTAAGCCTACCTGGGA 534

Qy       7316 AGGTGACCACATCCACCTTTAAAC 7339
              |||||||||||||||||||||||||
Db        535 AGGTGACCACATCCACCTTTAAAC 558


RESULT 377
US-10-301-480-49432
; Sequence 49432, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49432
;   LENGTH: 582
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-49432

  Query Match             5.7%;  Score 426.8;  DB 1;  Length 582;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  516;  Conservative    2;  Mismatches   59;  Indels   10;  Gaps    5;

Qy       1599 AACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG--AAAAAACTT 1656
              |||||||||||||   |||||||||||||||| |||||||||||||| ||  |||||||||
Db          1 AACTAATCTTAAAAGATAAGTTTATCACTCAGCCAGCTGCAGACATCAGAAAAAAAACTT 60

Qy       1657 CAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCG 1716
              |||||||||||  |||||||||||||||||||||||||||||| |||||||||||||:|||
Db         61 CAAAAGTCTGCTTTAGGCCCGGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCG 120

Qy       1717 GTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA 1776
              |||||||||||||||||||||||||||||||| ||||||||| | | || ||| |||||
Db        121 GTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAACAGGACAGATGAGA-TAAGAAAAA 179

Qy       1777 GGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAA 1836
              |||||||||||||||||| |||||    ||||  ||||||||||||||||||    ||||
Db        180 GGCCACCGCTTTAGTCATGGCCCTC---CAAGCTGACTTTGGAGGCTCTGGAACGTGGAA 236

Qy       1837 AAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTT 1896
              |  |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db        237 AGCCTGGGCAAATCGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTT 296

Qy       1897 TAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGG 1956
              |||| ||||||||| || ||||| | ||| ||||| | |||||||||||||||| ||||||
Db        297 TAAAGAAGATTGTCGAA-TAGAAATGAGCTGCCCCCTTGTCCATGCCCCTTATCTCAAGG 355

Qy       1957 GAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAA 2016
              |||||| |||||||||||||||||| || |  |||||| ||||| ||||||||||||||||
Db        356 GAATCATTGGAAGGCCCACTGCCCCAAGGAATGAAGGTCGTTTGAATCAGAAGCCACTAA 415

Qy       2017 CCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCAC 2076
              || |||||||||||||||||||||||||||| |||||||| |||  ||||| |||||||
Db        416 CCTGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCTGCCCATACCATCAC 475

Qy       2077 CCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGA 2133
              ||||||||||:|  |||||||||||||||||||||||||||| | |   ||||||||||
Db        476 CCTCACAGAGCYCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGGA 535

Qy       2134 CACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
              ||||||   || ||||| | ||||||| | ||||||   |||||||
Db        536 CACTGGCATGGCCTTCTCAATCTTACTTTCCTGTCCTATACAACTGT 582


RESULT 378
US-10-301-480-662841
; Sequence 662841, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 662841
;   LENGTH: 582
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-662841

  Query Match             5.7%;  Score 426.8;  DB 1;  Length 582;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  516;  Conservative    2;  Mismatches   59;  Indels   10;  Gaps    5;

Qy       1599 AACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG--AAAAAACTT 1656
              |||||||||||||   |||||||||||||||| |||||||||||||| ||  |||||||||
Db          1 AACTAATCTTAAAAGATAAGTTTATCACTCAGCCAGCTGCAGACATCAGAAAAAAAACTT 60

Qy       1657 CAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCG 1716
              |||||||||||  |||||||||||||||||||||||||||||| |||||||||||||:|||
Db         61 CAAAAGTCTGCTTTAGGCCCGGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCG 120

Qy       1717 GTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA 1776
              |||||||||||||||||||||||||||||||| ||||||||| | | || ||| |||||
Db        121 GTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAACAGGACAGATGAGA-TAAGAAAAA 179

Qy       1777 GGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAA 1836
              |||||||||||||||||| |||||    ||||  ||||||||||||||||||    ||||
Db        180 GGCCACCGCTTTAGTCATGGCCCTC---CAAGCTGACTTTGGAGGCTCTGGAACGTGGAA 236

Qy       1837 AAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTT 1896
              |  |||||||||| |||||||||||||||||||||||||||||| |||||||||||||||||
Db        237 AGCCTGGGCAAATCGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTT 296

Qy       1897 TAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGG 1956
              |||| ||||||||| || ||||| | ||| ||||| | |||||||||||||||| ||||||
Db        297 TAAAGAAGATTGTCGAA-TAGAAATGAGCTGCCCCCTTGTCCATGCCCCTTATCTCAAGG 355

Qy       1957 GAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAA 2016
              |||||| |||||||||||||||||| || |  |||||| ||||| ||||||||||||||||
Db        356 GAATCATTGGAAGGCCCACTGCCCCAAGGAATGAAGGTCGTTTGAATCAGAAGCCACTAA 415
```

```
Qy       2017 CCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCAC 2076
             || |||||||||||||||||||||||||||||| |||||||| |||  ||||| |||||||
Db        416 CCTGATGATCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCTGCCCATACCATCAC 475

Qy       2077 CCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGA 2133
             ||||||||||:|  |||||||||||||||||||||||||| | |   ||||||||||
Db        476 CCTCACAGAGCYCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGGA 535

Qy       2134 CACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
             ||||||   || ||||| | ||||||| | ||||||   ||||||||
Db        536 CACTGGCATGGCCTTCTCAATCTTACTTTCCTGTCCTATACAACTGT 582


RESULT 379
US-10-301-480-297640/c
; Sequence 297640, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 297640
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-297640

  Query Match             5.7%;  Score 425.7;  DB 1;  Length 530;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  489;  Conservative    4;  Mismatches   32;  Indels   13;  Gaps    4;

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        530 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCAGCCCAAAACCCT-- 473

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
               |||||||||| |||||||||||| ||||||:|||||||||||||||||||||||||||
Db        472 --ACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 415

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             |||||||||||||||||||||| |||||||||||||||| ||||||||||||  | |||
Db        414 ACAATGGAACAACTTCAGCACAGTAATAAACACCACTTCCATTTTAGTAGGACTCCCTGT 355

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||| | |||||||| || |||||| ||||||||||||||||||||||||||||
Db        354 TTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTTAGCAATACTTT 295

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             | | |||| |||||||||||| |||||||||| || |||||||||||  ||||||||||||
Db        294 AGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGAATAGTCTGCCT 235

Qy       6372 ACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGC 6426
             |||||||||||||     ||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGC 175

Qy       6427 TCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAA 6486
             |||||||||||||||||||||||||||||||||| |||| |||:|||||||||||||||
Db        174 TCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTGAA 116

Qy       6487 CAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT 6546
             |||||||||||||   ||  | || :|||||||||||||||||||||||||||||||||||
Db        115 CAAGATTTATACA---ATCACGTACYTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT 59

Qy       6547 TTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAA 6604
             |||||||| |||||||||||||||||   |||||||||||||||||| |||||||||
Db         58 TTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAA 1


RESULT 380
US-10-301-480-911049/c
; Sequence 911049, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911049
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-911049

  Query Match             5.7%;  Score 425.7;  DB 1;  Length 530;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  489;  Conservative    4;  Mismatches   32;  Indels   13;  Gaps    4;

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        530 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCAGCCCAAAACCCT-- 473

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
               |||||||||| |||||||||||| ||||||:|||||||||||||||||||||||||||
Db        472 --ACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 415

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             |||||||||||||||||||||| |||||||||||||||| ||||||||||||  | |||
Db        414 ACAATGGAACAACTTCAGCACAGTAATAAACACCACTTCCATTTTAGTAGGACTCCCTGT 355

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||| | |||||||| || |||||| ||||||||||||||||||||||||||||
Db        354 TTCCAATCTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTTAGCAATACTTT 295

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             | | |||| |||||||||||| |||||||||| || |||||||||||  ||||||||||||
Db        294 AGAGACAATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGAATAGTCTGCCT 235

Qy       6372 ACCCTCAGGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGC 6426
             |||||||||||||     ||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCCTCAGGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGC 175

Qy       6427 TCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAA 6486
             |||||||||||||||||||||||||||||||||| |||| |||:|||||||||||||||
Db        174 TCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTGAA 116

Qy       6487 CAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT 6546
             |||||||||||||   ||  | || :|||||||||||||||||||||||||||||||||||
Db        115 CAAGATTTATACA---ATCACGTACYTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT 59

Qy       6547 TTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAA 6604
             |||||||| |||||||||||||||||   |||||||||||||||||| |||||||||
Db         58 TTTGTTATTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAA 1


RESULT 381
US-10-301-480-473605/c
; Sequence 473605, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 473605
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-473605

  Query Match             5.7%;  Score 425.2;  DB 1;  Length 560;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  470;  Conservative   13;  Mismatches   61;  Indels   12;  Gaps    1;

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
             ||||||| :|||| || |||||| ||:| | ||:||||||| |||||||||| ||||||
Db        560 AGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGG 501

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
             |||   || ||:||||||||:|||||||||||||||||||| |||||||| |||||||||
Db        500 CCTTGCCATCCAATGGATGCCYTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAAT 441

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
             ||||||||||||||||||||||||||||||:|||||||||||| ||||||||||||||
Db        440 ATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 381

Qy       7017 AATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
             ||||  ||||||:|| |||            |||||||||||| ||||||| |||| |||
Db        380 AATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGAC 321

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
             ||| ||||||| |||| ||||||| ||||| :||||||| | ||||||||||||||| ||
Db        320 TAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATAT 261

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
               ||| |||||||||||||||||||| |||||||||| |||||| || ||||||||||
Db        260 GGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGC 201

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
             |||||||| |||||| | |:|| |||||||||||||||||||||| ||||||||| |||
Db        200 AGGAAGCATTTAGAGTGATCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTG 141

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
             ||| | || ||||||||||||  ||| ||||||||||||||||:|||||||||:||||
Db        140 AGAGGTGGAACTGAGAGACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAG 81

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
             |||||:||| |||||||| | |||||||||||||| |||||||||||||| ||:|||||
Db          80 CCTAGCTGGCAAGGTGACTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACAT 21

Qy       7365 CTGACCAATCAGAGAG 7380
             |  |||||  ||| ||
Db          20 CGAACCAAAAAGATAG 5


RESULT 382
US-10-301-480-1087014/c
; Sequence 1087014, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1087014
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1087014

  Query Match             5.7%;  Score 425.2;  DB 1;  Length 560;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  470;  Conservative   13;  Mismatches   61;  Indels   12;  Gaps    1;

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
             ||||||| :|||| || |||||| ||:| | ||:||||||| |||||||||| ||||||
Db        560 AGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGG 501

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
             |||   || ||:||||||||:|||||||||||||||||||| |||||||| |||||||||
Db        500 CCTTGCCATCCAATGGATGCCYTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAAT 441

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
             ||||||||||||||||||||||||||||||:|||||||||||| ||||||||||||||
Db        440 ATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 381

Qy       7017 AATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
             ||||  ||||||:|| |||            |||||||||||| ||||||| |||| |||
Db        380 AATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGAC 321

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
             ||| ||||||| |||| ||||||| ||||| :||||||| | ||||||||||||||| ||
Db        320 TAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATAT 261

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
               ||| |||||||||||||||||||| |||||||||| |||||| || ||||||||||
Db        260 GGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGC 201

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
             |||||||| |||||| | |:|| |||||||||||||||||||||| ||||||||| |||
Db        200 AGGAAGCATTTAGAGTGATCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTG 141

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
             ||| | || ||||||||||||  ||| ||||||||||||||||:|||||||||:||||
Db        140 AGAGGTGGAACTGAGAGACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAG 81

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
             |||||:||| |||||||| | |||||||||||||| |||||||||||||| ||:|||||
Db          80 CCTAGCTGGCAAGGTGACTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACAT 21

Qy       7365 CTGACCAATCAGAGAG 7380
             |  |||||  ||| ||
Db          20 CGAACCAAAAAGATAG 5


RESULT 383
US-10-301-480-414023/c
; Sequence 414023, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
```

```
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 414023
;    LENGTH: 563
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-414023

  Query Match             5.7%;  Score 424.6;  DB 1;  Length 563;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  501;  Conservative    2;  Mismatches   57;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| ||||||||||||||||||||||||||||||| ||||||
pb        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||
Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||| |||| |  || ||||||  |||||| : |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCYCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
             ||||||| |||||||||||| ||    |||||||||||||||||||||||||| |    |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
             ||||||| |||||||||||| | ||||| |||||||||| |||||| ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
             ||||||||||||||||   | | |||  | || ||| | ||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
             |||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||| | ||||||||||||||
Db          28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 384
US-10-301-480-414024/c
; Sequence 414024, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 414024
;    LENGTH: 563
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-414024

  Query Match             5.7%;  Score 424.6;  DB 1;  Length 563;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  501;  Conservative    2;  Mismatches   57;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| ||||||||||||||||||||||||||||||| ||||||
Db        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||
Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||| |||| | :|| ||||||  |||||||  |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTMCCTAGGGGATGAAGGTCCCCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
             ||||||| |||||||||||| ||    |||||||||||||||||||||||||| |    |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
             ||||||| |||||||||||| | ||||| |||||||||| |||||| ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
             ||||||||||||||||   | | |||  | || ||| | ||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
             |||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||| | ||||||||||||||
Db          28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 385
US-10-301-480-1027432/c
; Sequence 1027432, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1027432
;    LENGTH: 563
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-1027432

  Query Match             5.7%;  Score 424.6;  DB 1;  Length 563;
```

```
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  501;  Conservative    2;  Mismatches   57;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
              || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
              |||||||||||||||||||| |||||||||||||||||||||||||||||||| ||||||
Db        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
              |||||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||
Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
              |||| |||||||||||||||||| |||| |  || ||||||  |||||| : |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCYCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
              |||||||||| |||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
              ||||||| |||||||||||| ||    |||||||||||||||||||||||||| |    |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
              ||||||| |||||||||||| | ||||| |||||||||||| |||||| ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
              ||||||||||||||||    |  | |||  | || ||| | ||||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
              ||||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||| | ||||||||||||||
Db          28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 386
US-10-301-480-1027433/c
; Sequence 1027433, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1027433
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1027433

  Query Match             5.7%;  Score 424.6;  DB 1;  Length 563;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  501;  Conservative    2;  Mismatches   57;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
              || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||

Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
              |||||||||||||||||||| |||||||||||||||||||||||||||||||| ||||||

Db        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
              |||||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||

Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
              |||| |||||||||||||||||| |||| | :|| ||||||  ||||||   |||||||

Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTMCCTAGGGGATGAAGGTCCCCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
              |||||||||| |||||||||||||||||||||||||||||||||||||||| ||| ||

Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
              ||||||| |||||||||||| ||    |||||||||||||||||||||||||| |    |

Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
              ||||||| |||||||||||| | ||||| |||||||||||| |||||| ||||||||||||

Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
              ||||||||||||||||    |  | |||  | || ||| | ||||||||||||||| ||:||

Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
              ||||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||

Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||| | ||||||||||||||
Db          28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 387
US-10-029-386-6530
; Sequence 6530, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6530
;   LENGTH: 507
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP001697.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.91
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 2.00e-22
;   OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
US-10-029-386-6530

  Query Match             5.6%;  Score 423;  DB 1;  Length 507;
```

```
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  434;  Conservative   40;  Mismatches   21;  Indels    6;  Gaps    2;

Qy       3399 ACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAA 3458
             |||||||:|||||||||||||:|||||||  :|||||||||||||:||:|||::|||||| |
Db         11 ACCCAAGCCCCAGTGTTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATATGTCACAG-A 69

Qy       3459 AAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRT 3518
             ||||||||:||:|||||  :||||||||||| |||:|||:|||||:|||| ||||||||:|
Db         70 AAAAACAGGAATAGCTCCAGGAGTCCTTACGCAGATCCAAGGGACGAGCCTGCAACCCAT 129

Qy       3519 GGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGT 3578
             |||::||||||:||||||||:     |||||||||||||||:|:||||||||||:||||
Db        130 GGCATACCTGAGTAAGGAAAT-----TAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGT 184

Qy       3579 AGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCT 3638
             |||||  |||||  ||||::||||| ||||||||||||||| ||||| ||:|||||||||
Db        185 AGTGGCAGCAGTCACAGTCTTAGTAACTGAAGCAGTTAAAATGATACAAGGAAGAGATCT 244

Qy       3639 TACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCT 3698
             ||||||||||||||||||||:|||||::|||||||||||||||||||||||||||||| ||
Db        245 TACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGACT 304

Qy       3699 GTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACT 3758
             |||||||||:|||||||||||:|||||||||||||||||:|||||||| |||:|||
Db        305 GTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGTTGCGACT 364

Qy       3759 GTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARA 3818
             |||||||||||||||||||||||||:| |||| |||||||||||||||||||||||:|
Db        365 GTGCACTTGTGCAACTCTTAACCCAGCCACATTGCTTCCAGACAATGAAGAAAAGATAGA 424

Qy       3819 AYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGT 3878
             |:|||||||||||||:|||||  |||||||||||   |  |||||||||||||:||||:||
Db        425 ACATAACTGTCAACAAATAATTGCTCAAACCTACACTGCTCGAGGGGACCTTTTAGAAGT 484

Qy       3879 TCCYTTGACTGATCCYGACCT 3899
             |||:||||||||||||:|| ||
Db        485 TCCCTTGACTGATCCCGATCT 505


RESULT 388
US-10-301-480-414025/c
; Sequence 414025, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 414025
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-414025

  Query Match             5.6%;  Score 423;  DB 1;  Length 563;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  500;  Conservative    2;  Mismatches   58;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db         563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||

Db         503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||||| ||| ||||  ||||| || :|||||||||
Db         443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCAKCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||||| |||| |  || ||||||  ||||||   |||||||
Db         383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCCCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db         323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
             ||||||| |||||||||||| ||    |||||||||||||||||||||||| |    |
Db         263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
             ||||||| |||||||||||| | ||||| |||||||||||| |||||| ||||||||||||
Db         203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
             ||||||||||||||    | | |||  | || ||| | |||||||||||||| ||:||
Db         143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
             |||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||||
Db          87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             |||||||||||| | ||||||||||||||
Db          28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 389
US-10-301-480-414026/c
; Sequence 414026, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 414026
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-414026

  Query Match             5.6%;  Score 423;  DB 1;  Length 563;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  500;  Conservative    2;  Mismatches   58;  Indels    8;  Gaps    4;

Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db         563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||:

Db         503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTR 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||
Db         443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384
```

```
Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||| |||| |  || ||||||  ||||||   |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCCCTGAGTCA 324

Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
             |||||||  |||||||||||| ||    |||||||||||||||||||||||||||    |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
             |||||||  |||||||||||| | |||||  |||||||||||  ||||||  ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
             ||||||||||||||||    |  | |||  | || ||| | ||||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
             ||||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||  | ||||||||||||
Db         28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1
```

RESULT 390
US-10-301-480-1027434/c
; Sequence 1027434, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1027434
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1027434

Query Match 5.6%; Score 423; DB 1; Length 563;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 500; Conservative 2; Mismatches 58; Indels 8; Gaps 4;

```
Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||||| ||| ||||  ||||| || :|||||||||
Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCAKCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||| |||| |  || ||||||  ||||||   |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCCCTGAGTCA 324
```

```
Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
             |||||||  |||||||||||| ||    |||||||||||||||||||||||||||    |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
             |||||||  |||||||||||| | |||||  |||||||||||  ||||||  ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
             ||||||||||||||||    |  | |||  | || ||| | ||||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
             ||||||||||||||| || |||||||| ||||| ||||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
              |||||||||||  | ||||||||||||
Db         28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1
```

RESULT 391
US-10-301-480-1027435/c
; Sequence 1027435, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1027435
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1027435

Query Match 5.6%; Score 423; DB 1; Length 563;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 500; Conservative 2; Mismatches 58; Indels 8; Gaps 4;

```
Qy       1766 ATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCT 1825
             || |||||||| ||||| |||||| ||||| |||||||| ||| ||||||||||||||||
Db        563 ATAAAAAAAAAAGCCACTGCTTTAATCATGGCCCTCAGGGAAGCGGACTTTGGAGGCTCT 504

Qy       1826 GGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTA 1885
             |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        503 GGAAAAGGGAAAAGCTGGGCAAGTTGAATGCCTAATAGGGCTTGCTTCCAGTGTGGTCTA 444

Qy       1886 CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCC 1945
             |||||||||||||||||||||||||||||| ||| ||||  ||||| || ||||||||||
Db        443 CAAGGACACTTTAAAAAAGATTGTCCAAGTGGAAATAAGTTGCCCCCTCATCCATGCCCC 384

Qy       1946 TTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCA 2005
             |||| |||||||||||||||||| |||| |  || ||||||  ||||||   |||||||
Db        383 TTATGTCAAGGGAATCACTGGAAGACCCATTTACCTAGGGGATGAAGGTCCCCTGAGTCA 324
```

```
Qy       2006 GAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCC 2065
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        323 GAAGCCACTAATCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCACCAGCC 264

Qy       2066 CATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG---GT 2122
```

```
                |||||||  ||||||||||||  ||    |||||||||||||||||||||||||  |     |
Db        263 CATGCCAACACCCTCACAGATCCTCATGTATGCTTGACCATTGAGGGCCAGGAGGTTAAT 204

Qy       2123 TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCC 2182
                |||||||  ||||||||||||  | |||||  ||||||||||  ||||||  ||||||||||||
Db        203 TGTCTCCGGGACACTGGTGCAGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCC 144

Qy       2183 TCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTC 2242
                |||||||||||||||    |  | |||  | || |||  | ||||||||||||||| ||:||
Db        143 TCCAGATCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTC 88

Qy       2243 CCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGC 2302
                |||||||||||||| || ||||||||  |||||  ||||||||||||||||||||||||||
Db         87 CCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGC 29

Qy       2303 TTGAAAGCCCCACTACCTTGTTAGGGAG 2330
                ||||||||||| | ||||||||||||
Db         28 CTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 392
US-09-925-065A-406936/c
; Sequence 406936, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406936
;   LENGTH: 551
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-406936

  Query Match             5.6%;  Score 422;  DB 1;  Length 551;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  465;  Conservative   13;  Mismatches   58;  Indels   12;  Gaps    1;

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
                |||||||  :||||  ||  ||||||  ||:|  |  ||:|||||||  ||||||||||  ||||||
Db        551 AGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGG 492

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
                |||    ||  ||:|||||||||:|||||||||||||||||||  ||||||||  |||||||||
Db        491 CCTTGCCATCCAATGGATGCCYTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAAT 432

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
                |||||||||||||||||||||||||||||:|||||||||||||  ||||||||||||||
Db        431 ATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 372

Qy       7017 AATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
                ||||  ||||||:|| |||            ||||||||||| |||||||  |||| |||
Db        371 AATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGAC 312

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
                |||  |||||||  ||||  |||||||  |||||  :|||||||  |  ||||||||||||||  ||
Db        311 TAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATAT 252

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
                  |||  ||||||||||||||||||||||||  ||||||||||  ||||||  ||  ||||||||||
Db        251 GGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGC 192

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
                |||||||  ||||||  |  |:||  |||||||||||||||||||||||  |||||||||  |||
Db        191 AGGAAGCATTTAGAGTGATCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTG 132

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
                |||  |  ||  ||||||||||||  |||  |||||||||||||||||:|||||||||:||||
Db        131 AGAGGTGGAACTGAGAGACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAG 72

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
                |||||:|||  ||||||||  |  ||||||||||||||  ||||||||||||  ||:|||||
Db         71 CCTAGCTGGCAAGGTGACTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACAT 12

Qy       7365 CTGACCAA 7372
                |  |||||
Db         11 CGAACCAA 4


RESULT 393
US-09-925-065A-406936/c
; Sequence 406936, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406936
;   LENGTH: 551
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-406936

  Query Match             5.6%;  Score 422;  DB 1;  Length 551;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  465;  Conservative   13;  Mismatches   58;  Indels   12;  Gaps    1;

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
                |||||||  :||||  ||  ||||||  ||:|  |  ||:|||||||  ||||||||||  ||||||
Db        551 AGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGG 492

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
                |||    ||  ||:|||||||||:|||||||||||||||||||  ||||||||  |||||||||
Db        491 CCTTGCCATCCAATGGATGCCYTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAAT 432

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
                |||||||||||||||||||||||||||||:|||||||||||||  ||||||||||||||
Db        431 ATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 372

Qy       7017 AATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
                ||||  ||||||:|| |||            ||||||||||| |||||||  |||| |||
Db        371 AATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGAC 312

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
                |||  |||||||  ||||  |||||||  |||||  :|||||||  |  ||||||||||||||  ||
Db        311 TAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATAT 252
```

```
Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
              ||| |||||||||||||||||||||| |||||||||| |||||| || ||||||||||
Db        251 GGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGC 192

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
              |||||||| |||||| | |:|| |||||||||||||||||||||| ||||||||| |||
Db        191 AGGAAGCATTTAGAGTGATCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTG 132

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
              ||| | || |||||||||||   ||| ||||||||||||||||:||||||||:||||
Db        131 AGAGGTGGAACTGAGAGACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAG 72

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
              |||||:||| ||||||||  | |||||||||||||| |||||||||||||| ||:|||||
Db         71 CCTAGCTGGCAAGGTGACTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACAT 12

Qy       7365 CTGACCAA 7372
              |  |||||
Db         11 CGAACCAA 4


RESULT 394
US-09-925-065A-210293/c
; Sequence 210293, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210293
;   LENGTH: 572
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210293

  Query Match             5.6%;  Score 421.3;  DB 1;  Length 572;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  480;  Conservative   12;  Mismatches   59;  Indels   21;  Gaps    2;

Qy       6854 TACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGA 6913
              |||||| ||:| | ||:||||||| |||||||||| |||||||||   || ||:|||||
Db        572 TACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGGCCTTGCCATCCAATGGA 513

Qy       6914 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 6973
              |||| |||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db        512 TGCCTTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAATATTGCTACTCCTCTTTG 453

Qy       6974 GACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAAC 7033
              |||||||||||||:||||||||||| |||||||||||||||||  ||||||:|| |
Db        452 GACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAGAGCTGTAAAGC 393

Qy       7034 TA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 7081
              ||            |||||||||||| ||||||| |||| |||||| ||||||| ||||
Db        392 TACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGACTAAAATCTACCACAGAT 333

Qy       7082 CCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 7141
              ||||||| ||||| :|||||||| | |||||||||||||| ||  ||| ||||||||||||
```

```
Db        332 TCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATATGGAAGCCACCCCTCCTG 273

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
              |||||||||| |||||||||| |||||| || |||||||||||||||||. |||||| |
Db        272 AGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGCAGGAAGCATTTAGAGTG 213

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
              :|:|| |||||||||||||||||||||| |||||||| |||||| | || ||||||||
Db        212 RTCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTGAGAGGTGGAACTGAGAG 153

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
              ||||  ||| ||||||||||||||||||:|||||||||:|||||||||:||| |||||||
Db        152 ACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGCAAGGTGA 93

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA--------- 7372
              |  | |||||||||||||| |||||||||||| ||:||||| |  |||||
Db         92 CTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACATCGAACCAAAAAGATAGT 33

Qy       7373 TCAGAGAGCTCACTAAAATGCTAATTAGGCAA 7404
              ||| |||||||||||||||||||| ||||
Db         32 AAAGACAGCTCACTAAAATGCTAATTATGCAA 1


RESULT 395
US-09-925-065A-210293/c
; Sequence 210293, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 210293
;   LENGTH: 572
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-210293

  Query Match             5.6%;  Score 421.3;  DB 1;  Length 572;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  480;  Conservative   12;  Mismatches   59;  Indels   21;  Gaps    2;

Qy       6854 TACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGA 6913
              |||||| ||:| | ||:||||||| |||||||||| |||||||||   || ||:|||||
Db        572 TACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGGCCTTGCCATCCAATGGA 513

Qy       6914 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 6973
              |||| |||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db        512 TGCCTTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAATATTGCTACTCCTCTTTG 453

Qy       6974 GACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAAC 7033
              |||||||||||||:||||||||||| |||||||||||||||||  ||||||:|| |
Db        452 GACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAGAGCTGTAAAGC 393

Qy       7034 TA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 7081
              ||            |||||||||||| ||||||| |||| |||||| ||||||| ||||
Db        392 TACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGACTAAAATCTACCACAGAT 333

Qy       7082 CCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 7141
```

```
                 ||||||| ||||| :||||||| | |||||||||||||| ||  ||| ||||||||||
Db        332 TCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATATGGAAGCCACCCCTCCTG 273

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
                 |||||||||| |||||||||| |||||| || |||||||||||||||||| |||||| |
Db        272 AGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGCAGGAAGCATTTAGAGTG 213

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
                 :|:|| |||||||||||||||||||||||| ||||||||| |||||| | || ||||||||
Db        212 RTCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTGAGAGGTGGAACTGAGAG 153

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
                 ||||  ||| |||||||||||||||||:|||||||||:|||||||||:||| |||||||
Db        152 ACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGCAAGGTGA 93

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA--------- 7372
                 |  | |||||||||||||| |||||||||||||| ||:||||| |  |||||
Db         92 CTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACATCGAACCAAAAAGATAGT 33

Qy       7373 TCAGAGAGCTCACTAAAATGCTAATTAGGCAA 7404
                 ||| |||||||||||||||||||| ||||
Db         32 AAAGACAGCTCACTAAAATGCTAATTATGCAA 1
```

RESULT 396
US-09-925-065A-952783
; Sequence 952783, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 952783
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-952783

Query Match 5.6%; Score 421.1; DB 1; Length 549;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 486; Conservative 2; Mismatches 51; Indels 11; Gaps 3;

```
Qy       1668 CGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAA 1727
                 | ||||||| |||||||||||||||||||| |||||||||||||:|| ||||||||||
Db          1 CTTAGGCCCAGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCAGTTTTTTATAA 60

Qy       1728 TAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT-------TAAAAAAAAGGCC 1780
                 |||||||||||| ||||||| ||||  ||||||| |||||         ||||||||||
Db         61 TAGAGATCAGGAAGAGCAGG-GGAATGGGACAAATGGGATAAAAAAAAAAAAAAAAGGTG 119

Qy       1781 ACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGC 1840
                 || |||||||||:|| |||||||||||| ||||||||||||||||  ||||||||||||||
Db        120 ACTGCTTTAGTCRTGGCCCTCAGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGC 179

Qy       1841 TGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAA 1900
                 || |||||||||||||||| |||||||||||| |||| ||||||||||||||||||||||
Db        180 TGAGCAAATTGAATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAA 239
```

```
Qy       1901 AAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAAT 1960
                 |||||||||||||||||||| |||| ||||| | |||||||||||||| ||||||||||
Db        240 AAAGATTGTCCAAGTAGAAACAAGCTGCCCCCTTGTCCATGCCCCTTATGTCAAGGGAAT 299

Qy       1961 CACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAG 2020
                 |||||||||||||||||||||||| ||  |||||| | ||||||||||||||||||||||
Db        300 CACTGGAAGGCCCACTGCCCCAGGAGATGAAGGTCCTCTGAGTCAGAAGCCACTAACCAG 359

Qy       2021 ATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTC 2080
                 || |||||||||||||||||| ||||  |||||||||||| ||||||||||||||||||||
Db        360 ATAATCCAGCAGCAGGACTGAGGATGCCCAGGGCAAGCGCCAGCCCATGCCATCACCCTC 419

Qy       2081 ACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACT 2137
                 |||||||| |||||||||||||||||||||||||||||||| | |   |||||| ||||||||
Db        420 ACAGAGCCTTGGGTATGCTTGACCATTGAGGGCCAGGAGGTTCACTGTCTCTTGGACACT 479

Qy       2138 GGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACT 2197
                 |||  || ||||| |||||||||| |||||| |||||||||||||| ||||||||||||||
Db        480 GGTATGGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCCTTCAGATCTGTCACT 539

Qy       2198 ATTCTGAGGG 2207
                 || |   |||
Db        540 ATCCGAGGGG 549
```

RESULT 397
US-09-925-065A-952783
; Sequence 952783, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 952783
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-952783

Query Match 5.6%; Score 421.1; DB 1; Length 549;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 486; Conservative 2; Mismatches 51; Indels 11; Gaps 3;

```
Qy       1668 CGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAA 1727
                 | ||||||| |||||||||||||||||||| |||||||||||||:|| ||||||||||
Db          1 CTTAGGCCCAGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCAGTTTTTTATAA 60

Qy       1728 TAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT-------TAAAAAAAAGGCC 1780
                 |||||||||||| ||||||| ||||  ||||||| |||||         ||||||||||
Db         61 TAGAGATCAGGAAGAGCAGG-GGAATGGGACAAATGGGATAAAAAAAAAAAAAAAAGGTG 119

Qy       1781 ACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGC 1840
                 || |||||||||:|| |||||||||||| ||||||||||||||||  ||||||||||||||
Db        120 ACTGCTTTAGTCRTGGCCCTCAGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGC 179

Qy       1841 TGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAA 1900
                 || |||||||||||||||| |||||||||||| |||| ||||||||||||||||||||||
Db        180 TGAGCAAATTGAATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAA 239
```

```
Qy       1901 AAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAAT 1960
              ||||||||||||||||||  |||| ||||| | ||||||||||||||| ||||||||||
Db        240 AAAGATTGTCCAAGTAGAAACAAGCTGCCCCCTTGTCCATGCCCCTTATGTCAAGGGAAT 299

Qy       1961 CACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAG 2020
              |||||||||||||||||||||| ||  |||||| | ||||||||||||||||||||||
Db        300 CACTGGAAGGCCCACTGCCCCAGGAGATGAAGGTCCTCTGAGTCAGAAGCCACTAACCAG 359

Qy       2021 ATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTC 2080
              || ||||||||||||||||||| ||||  |||||||||||| ||||||||||||||||||
Db        360 ATAATCCAGCAGCAGGACTGAGGATGCCCAGGGCAAGCGCCAGCCCATGCCATCACCCTC 419

Qy       2081 ACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACT 2137
              |||||||| |||||||||||||||||||||||||||| | |   |||||| ||||||||
Db        420 ACAGAGCCTTGGGTATGCTTGACCATTGAGGGCCAGGAGGTTCACTGTCTCTTGGACACT 479

Qy       2138 GGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACT 2197
              |||  || ||||| |||||||||| |||||| |||||||||||| ||||||||||||
Db        480 GGTATGGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCCTTCAGATCTGTCACT 539

Qy       2198 ATTCTGAGGG 2207
              || |   |||
Db        540 ATCCGAGGGG 549
```

RESULT 398
US-09-864-761-16507
; Sequence 16507, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16507
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007567.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
US-09-864-761-16507

Query Match 5.6%; Score 420.9; DB 1; Length 558;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 512; Conservative 8; Mismatches 36; Indels 11; Gaps 7;

```
Qy        780 GTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACA 839
              | | | |||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 GAAAAGAAGGCAAATGGAGTGAAGTGCCATATGTACAAACTTTCTTTTCATTAAGAGACA 60

Qy        840 ACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTC 899
              |||| ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db         61 ACTCGCAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTCTCAGAGTCTACCTC 120

Qy        900 CCTATCCCAGCATCCCC--GACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
              |||| ||| |||||||  |||||||| |||::|||||| | ||||||||||||||||||
Db        121 CCTACCCCGGCATCCCCCTGACTCCTTCTCCAACTAATAAGCAACCCCCTTCAACCCAAA 180

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
              |||||||||||||||||||||||| ||||||| | ||||||||||||||||||||||||||
Db        181 TGGTCCAAAAGGAGATAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCC 240

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
                ||||| ||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 GTTTATGCCCCCT-CCAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTA 299

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
              |:||||::|||| ||||||| |||||||| |||| |||| ||||| ||||||||||||:
Db        300 CCTTTTTCTCTCTCAGACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAAC 359

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
              |||||||||||||||:|| |||||||||||||||||    | ||||||||||||||||
Db        360 CCTGATGGCTATATTGATG-TTTACAAGGGTTAGGACAA---TCTGATCTGACATGGAGA 415

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
              ||| || |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GATGTA-ATGTTACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 474

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
              ||||||| |||:|||||||||||||||||||||||||||||||||||||     ||||||||
Db        475 TGCAGCCCGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ACAGGATGACA 532

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCAC 1344
              |||| |||||||| || |||| |||||
Db        533 ACAG-AGGAAAGAGAACGATTGCCCAC 558
```

RESULT 399
US-10-029-386-4312/c
; Sequence 4312, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288

```
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4312
;   LENGTH: 521
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC021619.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
;   OTHER INFORMATION: EST_HUMAN HIT: BE734284.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 4.00e-26
;   OTHER INFORMATION: NT HIT: AF208161.1, EVALUE 0.00e+00
US-10-029-386-4312

  Query Match             5.6%;  Score 420.2;  DB 1;  Length 521;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  456;  Conservative   10;  Mismatches   43;  Indels   12;  Gaps    1;

Qy       6642 AAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 6701
             |||||| |||||||||| ||||| |||||||||||||||||||||||||||||||||||||
Db        521 AAATGGTGACATGGAATGGGTCACCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC 462

Qy       6702 CCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGG 6761
             | ||||||||||||||||||:||||||||||||||||||:|||||| || ||:|||||
Db        461 CTTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCCGAAAGAGG 402

Qy       6762 GGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGT 6821
             ||||| |||||||||||||| |||||||||||||||||| |||||||||| ||||||||
Db        401 GGGAATCTGTTTATTTTTAGTGGAAGAATGCTGTTATTACGTTAATCAATCTGGAATCGT 342

Qy       6822 CACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAA 6881
             ||| || ||||||:|||||||| |||||||||||||| ||:||||||:||||||| |||
Db        341 CACCGAAAAAGTTAAAGAAATTCAAGATCGAATACAACATAGAGCAGAGGAGCTTCAAAA 282

Qy       6882 CACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACC 6941
             |||  |||||||||| ||||||||||:||||||||||||||||||||||||||||||||
Db        281 CACCAGACCCTGGGACCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACC 222

Qy       6942 TCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAA 7001
             ||||||  |||||||| || ||||||||||||||||||||||:|||||||| ||||
Db        221 TCTAGCGACTATAATACTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTTTTAA 162

Qy       7002 CTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAA 7049
             ||||||||||||||||| |||||||:|| |||           |||||||||| |
Db        161 GTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTACAAATGGCTCTTCAAATGGAGCCCCA 102

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||| || |||| |||||| |||||| |  |||||||||||| |||||:|||  || | ||
Db        101 GATACAATCCATGACTAAAATCTACTGTGGACCCCTGGACCAGCCTGCTAGTGCATGCTC 42

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCT 7150
              |||||||||||||  |||||||||||||||||||||||||
Db         41 CCATGTTAATGACATTGAAGGCACCCCTCCTGAGGAAATCT 1


RESULT 400
US-09-925-065A-803434/c
; Sequence 803434, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803434
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-803434

  Query Match             5.6%;  Score 420.1;  DB 1;  Length 548;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  485;  Conservative   13;  Mismatches   47;  Indels    9;  Gaps    4;

Qy       6789 ATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG---AGAAAGTTRAAGAAATTCS 6845
             |||||||||||||||||||||||||||||| ||||||||||   |||||||:|| ||||||:
Db        548 ATGCTGTTATTATGTTAATCAATCCAGAATCGTCACTGAGAAGAAAGTTAAAAAAATTCG 489

Qy       6846 AGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAG 6905
             |||||| ||||||||||:||||| :||||||| |||||| |||||||||||||| ||||||
Db        488 AGATCGCATACAACGTAGAGCAGGGGAGCTTCAAAACACCGGACCCTGGGGCCCCCTCAG 429

Qy       6906 CCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACT 6965
             ||:|||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||
Db        428 CCAATGGATGCCCTGGATTCTCCCCTCCTTAGGACCTCTAGCAGCTATAATATTGTTACT 369

Qy       6966 CCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGC 7025
             |||||||||||||||||||| ||:|:|||||||||||| |||||||||||||||||| ||||
Db        368 CCTCTTTGGACCCTGTATC-TTAASCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAGC 310

Qy       7026 TGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCT 7085
             |||:|| |||||||||| |   | ||| || || |||   ||    ||   |||||||
Db        309 TGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAGT----CCATGGACCCCT 254

Qy       7086 GGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGA 7145
             ||||||||||:||||| | | |||||||||||||||||| ||||||  |||||||| |
Db        253 GGACCGGCCTGCTAGCCTATGCTCTGATGTTAATGACATCGAAGGCAT-CCTCCTGAGAA 195

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             ||||||| |||||||| ||||||| ||||| |||||||||||||||||||||||| |||:|
Db        194 AATCTCAACTGCACAACCGCTACTACACCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCG 135

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
               |||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db        134 CTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAG 75

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||| ||  |:||||||||:|||||||| :|||||||||||||||
Db         74 GACTAGCTGGATTTCCTAAGCCAACTAAGAATCCCTAAGCCTACCTGGGAAGGTGACCAC 15

Qy       7326 ATCCACCTTTAAAC 7339
              ||||||||||||||
Db         14 ATCCACCTTTAAAC 1


RESULT 401
US-09-925-065A-803434/c
; Sequence 803434, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803434
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-803434

  Query Match             5.6%;  Score 420.1;  DB 1;  Length 548;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  485;  Conservative   13;  Mismatches   47;  Indels    9;  Gaps    4;

Qy       6789 ATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG---AGAAAGTTRAAGAAATTCS 6845
             |||||||||||||||||||||||||| |||||||||||   ||||||||:|| ||||||:
Db        548 ATGCTGTTATTATGTTAATCAATCCAGAATCGTCACTGAGAAGAAAGTTAAAAAAATTCG 489

Qy       6846 AGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAG 6905
             |||||| ||||||||||:||||| :||||||| |||||| |||||||||||||| ||||||
Db        488 AGATCGCATACAACGTAGAGCAGGGGAGCTTCAAAACACCGGACCCTGGGGCCCCCTCAG 429

Qy       6906 CCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACT 6965
             ||:|||||||||||||||||||||| ||||||||||||||||||||||||||||||| ||||
Db        428 CCAATGGATGCCCTGGATTCTCCCCTCCTTAGGACCTCTAGCAGCTATAATATTGTTACT 369

Qy       6966 CCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGC 7025
             |||||||||||||||||| ||:|:|||||||||||| ||||||||||||||||||| ||||
Db        368 CCTCTTTGGACCCTGTATC-TTAASCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAGC 310

Qy       7026 TGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCT 7085
             |||:|| |||||||||| |   | ||| || || |||   ||      ||   |||||||
Db        309 TGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAGT----CCATGGACCCCT 254

Qy       7086 GGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGA 7145
             ||||||||||:||||| | | |||||||||||||||||| ||||||  |||||||| |
Db        253 GGACCGGCCTGCTAGCCTATGCTCTGATGTTAATGACATCGAAGGCAT-CCTCCTGAGAA 195

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             ||||||| |||||||||| ||||||| ||||| |||||||||||||||||||| |||:|
Db        194 AATCTCAACTGCACAACCGCTACTACACCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCG 135

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
               |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||
Db        134 CTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAG 75

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||| ||  |:|||||||||:|||||||| :|||||||||||||||
Db          74 GACTAGCTGGATTTCCTAAGCCAACTAAGAATCCCTAAGCCTACCTGGGAAGGTGACCAC 15

Qy       7326 ATCCACCTTTAAAC 7339
             ||||||||||||||
Db          14 ATCCACCTTTAAAC 1


RESULT 402
US-09-925-065A-956897
; Sequence 956897, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956897
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956897

  Query Match             5.6%;  Score 420.1;  DB 1;  Length 548;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  485;  Conservative    2;  Mismatches   51;  Indels   11;  Gaps    3;

Qy       1668 CGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAA 1727
             | ||||||| |||||||||||||||||||| ||||||||||||:|| ||||||||||
Db          1 CTTAGGCCCAGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCAGTTTTTTATAA 60

Qy       1728 TAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT-------TAAAAAAAAGGCC 1780
             |||||||||||| ||||||| |||| ||||||| |||||        ||||||||||
Db         61 TAGAGATCAGGAAGAGCAGG-GGAATGGGACAAATGGGATAAAAAAAAAAAAAAAAGGTG 119

Qy       1781 ACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGC 1840
             || |||||||||:|| |||||||||| ||||||||||||||  ||||||||||||||
Db         120 ACTGCTTTAGTCRTGGCCCTCAGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGC 179

Qy       1841 TGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAA 1900
             || |||||||||||||||| |||||||||| |||| ||||||||||||||||||||
Db         180 TGAGCAAATTGAATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAA 239

Qy       1901 AAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAAT 1960
             |||||||||||||||||||  |||| ||||| | ||||||||||||| ||||||||||
Db         240 AAAGATTGTCCAAGTAGAAACAAGCTGCCCCCTTGTCCATGCCCCTTATGTCAAGGGAAT 299

Qy       1961 CACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAG 2020
             |||||||||||||||||||||||| ||  |||||| | ||||||||||||||||||||||
Db         300 CACTGGAAGGCCCACTGCCCCAGGAGATGAAGGTCCTCTGAGTCAGAAGCCACTAACCAG 359

Qy       2021 ATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTC 2080
             || |||||||||||||||||| ||||  |||||||||||| ||||||||||||||||||
Db         360 ATAATCCAGCAGCAGGACTGAGGATGCCCAGGGCAAGCGCCAGCCCATGCCATCACCCTC 419

Qy       2081 ACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACT 2137
             |||||||| |||||||||||||||||||||||||||||||||| | |   |||||| ||||||||
Db         420 ACAGAGCCTTGGGTATGCTTGACCATTGAGGGCCAGGAGGTTCACTGTCTCTTGGACACT 479

Qy       2138 GGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACT 2197
             |||  || ||||| |||||||||||| |||||| |||||||||||||| ||||||||||||||
Db         480 GGTATGGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCCTTCAGATCTGTCACT 539

Qy       2198 ATTCTGAGG 2206
             || |   ||
Db         540 ATCCGAGGG 548


RESULT 403
US-09-925-065A-956897
; Sequence 956897, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 956897
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-956897

  Query Match             5.6%;  Score 420.1;  DB 1;  Length 548;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  485;  Conservative    2;  Mismatches   51;  Indels   11;  Gaps    3;

Qy       1668 CGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAA 1727
             | ||||||| |||||||||||||||||||||| |||||||||||||:|| ||||||||||
Db             1 CTTAGGCCCAGAGCAAAACTTAGAAACCCTACTGAACTTGGCAACCTCAGTTTTTTATAA 60

Qy       1728 TAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT-------TAAAAAAAAGGCC 1780
             |||·|||||||||| ||||||| |||| ||||||| |||||         ||||||||||
Db            61 TAGAGATCAGGAAGAGCAGG-GGAATGGGACAAATGGGATAAAAAAAAAAAAAAAAGGTG 119

Qy       1781 ACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGC 1840
             || |||||||||:|| |||||||||| |||||||||||||||  ||||||||||||||
Db            120 ACTGCTTTAGTCRTGGCCCTCAGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGC 179

Qy       1841 TGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAA 1900
             || |||||||||||||||||| |||||||||||| |||| ||||||||||||||||||||
Db            180 TGAGCAAATTGAATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAA 239

Qy       1901 AAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAAT 1960
             |||||||||||||||||||  |||| ||||| | |||||||||||||||| ||||||||||
Db            240 AAAGATTGTCCAAGTAGAAACAAGCTGCCCCCTTGTCCATGCCCCTTATGTCAAGGGAAT 299

Qy       1961 CACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAG 2020
             |||||||||||||||||||||||| ||  |||||| | ||||||||||||||||||||||
Db            300 CACTGGAAGGCCCACTGCCCCAGGAGATGAAGGTCCTCTGAGTCAGAAGCCACTAACCAG 359

Qy       2021 ATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTC 2080
             || |||||||||||||||||||| ||||  |||||||||||| ||||||||||||||||||
Db            360 ATAATCCAGCAGCAGGACTGAGGATGCCCAGGGCAAGCGCCAGCCCATGCCATCACCCTC 419

Qy       2081 ACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACT 2137
             |||||||| |||||||||||||||||||||||||||| | |   |||||| ||||||||
Db            420 ACAGAGCCTTGGGTATGCTTGACCATTGAGGGCCAGGAGGTTCACTGTCTCTTGGACACT 479

Qy       2138 GGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACT 2197
             |||  || |||||| |||||||||||| |||||| |||||||||||||| ||||||||||||
Db            480 GGTATGGCCTTCTCAGTCTTACTCTCCTGTCCTGGACAACTGTCCTTCAGATCTGTCACT 539

Qy       2198 ATTCTGAGG 2206
             || |   ||
Db            540 ATCCGAGGG 548


RESULT 404
US-09-925-065A-932905/c
; Sequence 932905, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932905
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932905

  Query Match             5.6%;  Score 419.9;  DB 1;  Length 565;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  503;  Conservative    8;  Mismatches   54;  Indels   11;  Gaps    5;

Qy       2371 ATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGT 2430
             |||||||||||||||| |||||| ||| ||||||||||||||||||||||||||||||||
Db             565 ATAGGAGAAGGAACACACGTTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGT 507

Qy       2431 CTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAAC 2490
             |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||
Db             506 CTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCCATCCTGTTCAAGTTAAAC 448

Qy       2491 TAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAAC 2550
             |||||||||| | |||||||||| |||| |||||||||||| |||||| || ||||||
Db             447 TAAAGGATTCCGCCTCCTTTCCCTAACAAAAGCAGTACCCCCTTAGACCCGAGCCCCAAC 388

Qy       2551 AAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAA 2610
             || || |||||||||| |||||||| |||||||||||||| |||||||| ||||||  |
Db             387 AATGACTCCAAAAGATTTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATA 328

Qy       2611 CTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGG 2670
               ||||||| || | |       | |   |  | |||:|||||||   |||||||| |
Db             327 GCCCCTGCAATACTCC-------AAATTTAGGAGTACARAAACCCATCAGACAGTGGA-G 276

Qy       2671 GTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTAC 2730
             |||||||||||||||||||||||||||| | |||||||||||  |  ||||||||||||||
Db             275 GTTAGTGCAAGATCTCAGGATTATCAGT-GAGGCCGTTGTCACTCCATACCCAGCTGTAC 217

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||| |||||||||| ||::|||||||||||||||||  |||||||||||:||||||
Db             216 CTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAATAGAGTGGTTTACAGTCCTGG 157

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             || ||:||| |||||| |||||||| |||| |||||||||||||| ||||||||||||||
Db             156 ACTTTAAGGCTGCCTTTTTCTGCATCGCTGTGCATCCTGACTCTCAGTTCTTGTTTGCCT 97

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             |||||||| ||||||| |||:  |||||||||||||||||:|||||||||||||||||||
Db              96 TTGAAGATCCTTCAAATCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCA 37

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCC 2946
              ||||||:|||||||||||||||||||||||||||||
Db              36 AGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCC 1


RESULT 405
US-09-925-065A-932905/c
; Sequence 932905, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932905
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932905

  Query Match             5.6%;  Score 419.9;  DB 1;  Length 565;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  503;  Conservative    8;  Mismatches   54;  Indels   11;  Gaps    5;

Qy       2371 ATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGT 2430
             |||||||||||||||| |||||| ||| |||||||||||||||||||||||||||||||||
Db        565 ATAGGAGAAGGAACACACGTTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGT 507

Qy       2431 CTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAAC 2490
             ||||||||||||||||||||||||||||| |||||||||| |||||||||||||||||||
Db        506 CTGGGCAACAGAAGGACAATATGGACGAG-CAAAGAATGCCCATCCTGTTCAAGTTAAAC 448

Qy       2491 TAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAAC 2550
             |||||||||| | |||||||||| |||| |||||||||||| |||||| || ||||||
Db        447 TAAAGGATTCCGCCTCCTTTCCCTAACAAAAGCAGTACCCCCTTAGACCCGAGCCCCAAC 388

Qy       2551 AAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAA 2610
             || || |||||||||||| |||||||| |||||||||||||| |||||||| ||||||  |
Db        387 AATGACTCCAAAAGATTTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATA 328

Qy       2611 CTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGG 2670
              ||||||| || | |         | |    |  | |||:||||||||   |||||||||  |
Db        327 GCCCCTGCAATACTCC-------AAATTTAGGAGTACARAAACCCATCAGACAGTGGA-G 276

Qy       2671 GTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTAC 2730
             |||||||||||||||||||||||||| | |||||||||||   |  ||||||||||||||
Db        275 GTTAGTGCAAGATCTCAGGATTATCAGT-GAGGCCGTTGTCACTCCATACCCAGCTGTAC 217

Qy       2731 CTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGG 2790
             ||| |||||||||| ||::||||||||||||||||||  |||||||||||||:||||||
Db        216 CTAACCCTTATACTCTGCTTTCCCAAATACCAGAGGAAATAGAGTGGTTTACAGTCCTGG 157

Qy       2791 ACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCT 2850
             || ||:||| |||||| |||||||| |||| |||||||||||||| ||||||||||||||
Db        156 ACTTTAAGGCTGCCTTTTTCTGCATCGCTGTGCATCCTGACTCTCAGTTCTTGTTTGCCT 97

Qy       2851 TTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCA 2910
             |||||||| ||||||| |||:  |||||||||||||||||:||||||||||||||||||
Db         96 TTGAAGATCCTTCAAATCCAATGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCA 37

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCC 2946
              |||||||||||||||||||||||||||||||
Db         36 AGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCC 1


RESULT 406
US-09-925-065A-406907/c
; Sequence 406907, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406907
;   LENGTH: 537
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-406907

  Query Match             5.6%;  Score 419.3;  DB 1;  Length 537;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  442;  Conservative   46;  Mismatches   48;  Indels    3;  Gaps    2;

Qy       3247 YAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACA 3306
             :|||:|||||| |||:||||||| |||||||:  ||:|||||||||  :|| ||:|||||
Db        537 CAGGTTTCTGCTGAATATGGATT-CCCAGGTACCGCAAAATAGCCACACCACTATATACA 479

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
             :|||||| | |||||||| |||||||||||||||:|||| ||||:| ||||||:::||
Db        478 CTAATTAAAGGAACTCAGACAGCCAATACCCATTTAGTAAGCTGGACATCTGAAGCAGAA 419

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
             | |||||||||||| |  ||  | | |||  ||| |||:||||||||||||:|| ||||
Db        418 GAGGCTTTCCAGGCACTAAAGAAGGCCCT--AACTTAAGCCCCAGTGTTAAGCTTCCCAA 361

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
             |:||:|||||||||||:||:|||::|||||||||||||| :||:||||| :||||||||||
Db        360 CAGGRCAAGACTTTTCTTTATATGTCACAGAAAAAAACAAGAATAGCTCCAGGAGTCCTT 301

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
             ||| ||:|| :|| ||:|||||||| ||||::|||||::|||||:||||||||:||||| |
Db        300 ACATAGGTCTGAGAGATGAGCTTGGAACCCATGGCATACCTGAGTAAGGAAATTGATGCA 241

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
             |||||||||||||| |:|||||||||:|| ||||||  |||||| || |::  ||||||
Db        240 GTGGCAAAGGGTTGTCCTCATTGTTTATGGATAGTGGCAGCAGTAACAATCTAGGTATCT 181

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              |||||||||||||||||| ||:|| |||||||||||||||||||||||||||||:|||||:
Db        180 TTAGCAGTTAAAATAATACAAGGAAGGGATCTTACTGTGTGGACATCTCATGATGTGAAC 121

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
             :|||||||||||||||||||||||||||||||||||||||||||||: ||||||||||:||||
Db        120 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTTTTTACTTAAATATCAG 61

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGY 3785
             ||||||||| ||||||: || ||||||| :| || |||||||||||||||||||||||||:
Db         60 GCTCTATTATTTGAAGTGCTAGTGCTGTGAATGCGCACTTGTGCAACTCTTAACCCAGC 2


RESULT 407
US-09-925-065A-406907/c
; Sequence 406907, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406907
;    LENGTH: 537
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-406907

  Query Match             5.6%;  Score 419.3;  DB 1;  Length 537;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  442;  Conservative   46;  Mismatches   48;  Indels    3;  Gaps    2;

Qy       3247 YAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACA 3306
              :|||:|||||| |||:||||||| |||||||:  ||:|||||||| :|| ||:|||||
Db        537 CAGGTTTCTGCTGAATATGGATT-CCCAGGTACCGCAAAATAGCCACACCACTATATACA 479

Qy       3307 STAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAA 3366
              :||||||| | |||||||| ||||||||||||||||:|||| ||||:| ||||||:::||
Db        478 CTAATTAAAGGAACTCAGACAGCCAATACCCATTTAGTAAGCTGGACATCTGAAGCAGAA 419

Qy       3367 GTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAA 3426
              | |||||||||||| |  ||  | | |||  |||  |||:||||||||||||:|| ||||
Db        418 GAGGCTTTCCAGGCACTAAAGAAGGCCCT--AACTTAAGCCCCAGTGTTAAGCTTCCCAA 361

Qy       3427 CRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTT 3486
              |:||:||||||||||:||:|||::|||||||||||||| :||:||||| :|||||||||
Db        360 CAGGRCAAGACTTTTCTTTATATGTCACAGAAAAAAACAAGAATAGCTCCAGGAGTCCTT 301

Qy       3487 ACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTA 3546
              ||| ||:|| :|| ||:||||||| ||||::||||::||||||:||||||||:||||| |
Db        300 ACATAGGTCTGAGAGATGAGCTTGGAACCCATGGCATACCTGAGTAAGGAAATTGATGCA 241

Qy       3547 GTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCT 3606
              |||||||||||||| |:||||||||:|| ||||||  |||||| || |::  ||||||
Db        240 GTGGCAAAGGGTTGTCCTCATTGTTTATGGATAGTGGCAGCAGTAACAATCTAGGTATCT 181

Qy       3607 GAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAY 3666
              |||||||||||||||||| ||:|| |||||||||||||||||||||||||||||:|||||:
Db        180 TTAGCAGTTAAAATAATACAAGGAAGGGATCTTACTGTGTGGACATCTCATGATGTGAAC 121

Qy       3667 RGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAG 3726
              :|||||||||||||||||||||||||||||||||||||||||: |||||||||||:||||
Db        120 AGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTTTTTACTTAAATATCAG 61

Qy       3727 GCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGY 3785
              ||||:|||| |||||: || ||||||| :| || |||||||||||||||||||||||||:
Db          60 GCTCTATTATTTGAAGTGCTAGTGCTGTGAATGCGCACTTGTGCAACTCTTAACCCAGC 2


RESULT 408
US-10-301-480-297638/c
; Sequence 297638, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 297638
;    LENGTH: 566
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-297638

  Query Match             5.6%;  Score 416.9;  DB 1;  Length 566;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  475;  Conservative   12;  Mismatches   58;  Indels   21;  Gaps    2;

Qy       6854 TACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGA 6913
              |||||| ||:| | ||:|||||||| |||||||||| |||||||||   || ||:|||||
Db        566 TACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGGCCTTGCCATCCAATGGA 507

Qy       6914 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 6973
              |||| |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db        506 TGCCTTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAATATTGCTACTCCTCTTTG 447

Qy       6974 GACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAAC 7033
              |||||||||||||:|||||||||||||| |||||||||||||||||  ||||||:|| |
Db        446 GACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAGAGCTGTAAAGC 387

Qy       7034 TA----------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 7081
              ||          |||||||||||| |||||| |||| |||||| |||||| ||||
Db        386 TACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGACTAAAATCTACCACAGAT 327

Qy       7082 CCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 7141
              ||||||| ||||| :|||||||| | ||||||||||||| ||  ||| ||||||||||
Db        326 TCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATATGGAAGCCACCCCTCCTG 267

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
              |||||||||| |||||||||| |||||| || |||||||||||||||||| |||||| |
Db        266 AGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGCAGGAAGCATTTAGAGTG 207

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
              :|:|| |||||||||||||||||||||| |||||||| |||||| | || ||||||||
Db        206 RTCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTGAGAGGTGGAACTGAGAG 147

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
              ||||  ||| |||||||||||||||||:|||||||||:|||||||||:||| |||||||
Db        146 ACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGCAAGGTGA 87

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA--------- 7372
              |  | |||||||||||||| |||||||||||||| ||:||||| |  |||||
Db          86 CTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACATCGAACCAAAAAGATAGT 27

Qy       7373 TCAGAGAGCTCACTAAAATGCTAATT 7398
              ||| ||||||||||||||||||||
Db          26 AAAGACAGCTCACTAAAATGCTAATT 1


RESULT 409
US-10-301-480-911047/c
; Sequence 911047, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911047
;    LENGTH: 566
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-911047
```

```
  Query Match             5.6%;  Score 416.9;  DB 1;  Length 566;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  475;  Conservative   12;  Mismatches   58;  Indels   21;  Gaps    2;

Qy       6854 TACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGA 6913
             |||||| ||:| | ||:||||||| |||||||||| |||||||||    || ||:|||||
Db        566 TACAACATAGAACCGAGGAGCTTCAAAACACTGGACTCTGGGGCCTTGCCATCCAATGGA 507

Qy       6914 TGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTG 6973
             |||| ||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db        506 TGCCTTGGATTCTCCCCTTCTTAGAACCTCTAGTAGCTATAATATTGCTACTCCTCTTTG 447

Qy       6974 GACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAAC 7033
             |||||||||||||:|||||||||||| ||||||||||||||||  ||||||:|| |
Db        446 GACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAGAGCTGTAAAGC 387

Qy       7034 TA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 7081
             ||            |||||||||||| ||||||| |||| |||||| ||||||| ||||
Db        386 TACAAATCATTCTTCAAATGGAGCCCCAGATGCAATCCATGACTAAAATCTACCACAGAT 327

Qy       7082 CCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 7141
              ||||||| ||||| :||||||| | |||||||||||||| ||  ||| ||||||||||
Db        326 TCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTAATGATATGGAAGCCACCCCTCCTG 267

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
             |||||||||||| |||||||||| |||||| || |||||||||||||||||| |||||| |
Db        266 AGGAAATCTCAACTGCACAACCCCTACTATGCTCCAATTCAGCAGGAAGCATTTAGAGTG 207

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
             :|:|| |||||||||||||||||||||| ||||||||| |||||| | || ||||||||
Db        206 RTCGTTGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTCTTGAGAGGTGGAACTGAGAG 147

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
             ||||  ||| ||||||||||||||||:|||||||||:|||||||||:||| |||||||
Db        146 ACAGAGCTACCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGCAAGGTGA 87

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAA--------- 7372
             |  | |||||||||||||| |||||||||||| ||:||||| |  |||||
Db          86 CTGCTTCCACCTTTAAACATGGGGCTTGCAACTGAGCTCACATCGAACCAAAAAGATAGT 27

Qy       7373 TCAGAGAGCTCACTAAAATGCTAATT 7398
               ||| ||||||||||||||||||||
Db          26 AAAGACAGCTCACTAAAATGCTAATT 1


RESULT 410
US-09-864-761-1777
; Sequence 1777, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1777
;  LENGTH: 488
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: MAP TO AC004897.1
;  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  10
;  OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
;  OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
;  OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
;  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
;  OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99
;  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
;  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
US-09-864-761-1777

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             |||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             |||||||||||  |||||||||||||||||||||  ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | |||||||||  |||||||||||||||||||||||| |||||||  |||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             ||||| |||||||| ||||||| |||||||||||  ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             |||| || | |||||||||||||||||||||||  | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419
```

```
Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             ||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
             | ||||||
Db        480 CTCTGAGTC 488


RESULT 411
US-09-925-065A-924272
; Sequence 924272, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 924272
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-924272

  Query Match             5.5%;  Score 409.4;  DB 1;  Length 564;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  498;  Conservative    3;  Mismatches   60;  Indels   10;  Gaps    5;

Qy       1710 AACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTA 1769
             | |:|   ||||||||| |||||||||||||||||||| ||||| ||||||||  || ||
Db          1 ACCTTATTTTTTTTATACTAGAGATCAGGAGGAGCAGGTGGAACGGGACAAACAAGA-TA 59

Qy       1770 AAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAA 1829
             | |||||||||||||||||||||||| |||||||| ||| ||||||||||||||||||||
Db         60 AGAAAAAGGCCACCGCTTTAGTCATGGCCCTCAGGAAAGCAGACTTTGGAGGCTCTGGAA 119

Qy       1830 AAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAG 1889
             | ||||| |||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        120 TATGGAAAGGCTGGGCAAATTGAATGCCTCATAGGRCTTGCTTCCAGTGCGGTCTACAAG 179

Qy       1890 GACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTAT 1949
             |||| |||||||||||||||| ||||| |||||||||| ||||||||||||||||||||
Db        180 GACATTTTAAAAAAGATTGTCCAAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTAT 239

Qy       1950 TTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAG 2009
             |||||||||||||||||||||| |||||||||||||  |||||| | ||||||||||||
Db        240 GTCAAGGGAATCACTGGAAGGCCAACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAG 299

Qy       2010 CCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATG 2069
             |||||||||||||||||||||||| ||||| |||||  ||||||||  ||| ||||||
Db        300 CCACTAACCAGATGATCCAGCAGCAGAACTGAAGGTGCTGGGGGCAAGTCCCAGCCCATG 359

Qy       2070 CCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTC 2126
             |||||||||||||||||  ||||||||||||||||  |||||||||  |   ||||
Db        360 CCATCACCCTCACAGAGCCCCAGGTATGCTTGACCATTGAAAGCCAGGAAGTTAACTGTC 419

Qy       2127 TCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCA 2186
             |||||||||||||| || | ||||| ||||||||| | |||||| ||||||||||||||||
Db        420 TCCTGGACACTGGCGCAGCCTTCTCAGTCTTACTTTCCTGTCCTGGACAACTGTCCTCCA 479

Qy       2187 GATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAG 2246
             || |||||||||   |  | |||  | || ||| | ||||||||||||| |||:||||||
Db        480 GACCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATA-TTTCTCCCAG 535

Qy       2247 CCACTAAGTTATGAACTGGGGAGCTTTATTC 2277
             |||||||||| |||  |||||| ||||| ||
Db        536 CCACTAAGTTGTGA--TGGGGAACTTTACTC 564


RESULT 412
US-09-925-065A-924272
; Sequence 924272, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 924272
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-924272

  Query Match             5.5%;  Score 409.4;  DB 1;  Length 564;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  498;  Conservative    3;  Mismatches   60;  Indels   10;  Gaps    5;

Qy       1710 AACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTA 1769
             | |:|   ||||||||| |||||||||||||||||||| ||||| ||||||||  || ||
Db          1 ACCTTATTTTTTTTATACTAGAGATCAGGAGGAGCAGGTGGAACGGGACAAACAAGA-TA 59

Qy       1770 AAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAA 1829
             | |||||||||||||||||||||||| |||||||| ||| ||||||||||||||||||||
Db         60 AGAAAAAGGCCACCGCTTTAGTCATGGCCCTCAGGAAAGCAGACTTTGGAGGCTCTGGAA 119

Qy       1830 AAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAG 1889
             | ||||| |||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        120 TATGGAAAGGCTGGGCAAATTGAATGCCTCATAGGRCTTGCTTCCAGTGCGGTCTACAAG 179

Qy       1890 GACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTAT 1949
             |||| |||||||||||||||| ||||| |||||||||| ||||||||||||||||||||
Db        180 GACATTTTAAAAAAGATTGTCCAAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTAT 239

Qy       1950 TTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAG 2009
             |||||||||||||||||||||| |||||||||||||  |||||| | ||||||||||||
Db        240 GTCAAGGGAATCACTGGAAGGCCAACTGCCCCAGGGGATGAAGGTCCTCTGAGTCAGAAG 299

Qy       2010 CCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATG 2069
             |||||||||||||||||||||||| ||||| |||||  ||||||||  ||| ||||||
Db        300 CCACTAACCAGATGATCCAGCAGCAGAACTGAAGGTGCTGGGGGCAAGTCCCAGCCCATG 359

Qy       2070 CCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTC 2126
             |||||||||||||||||  ||||||||||||||||  |||||||||  |   ||||
Db        360 CCATCACCCTCACAGAGCCCCAGGTATGCTTGACCATTGAAAGCCAGGAAGTTAACTGTC 419

Qy       2127 TCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCA 2186
```

```
                ||||||||||||| || | ||||| ||||||||| | |||||| ||||||||||||||||
Db        420 TCCTGGACACTGGCGCAGCCTTCTCAGTCTTACTTTCCTGTCCTGGACAACTGTCCTCCA 479

Qy       2187 GATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAG 2246
                || |||||||||   |  | |||  | || ||| | |||||||||||| |||:||||||
Db        480 GACCTGTCACTA---TCTGAGGGGTCCTAGGACAGCCAGTCACTAGATA-TTTCTCCCAG 535

Qy       2247 CCACTAAGTTATGAACTGGGGAGCTTTATTC 2277
                |||||||||| |||  |||||| ||||| ||
Db        536 CCACTAAGTTGTGA--TGGGGAACTTTACTC 564


RESULT 413
US-09-925-065A-926278
; Sequence 926278, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926278
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926278

  Query Match             5.4%;  Score 404.1;  DB 1;  Length 552;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  476;  Conservative   26;  Mismatches   42;  Indels   27;  Gaps    5;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
                |||||||:|||:| ||||||||||||    |||||:||||| ||:|| ||||||||||||
Db          1 AATAGCCCTCACTCGGGCACTAGAATT---AGAAGGAAAAAAGGTAAGTATATATACAGA 57

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
                ||||:|:|||||||:|||||||| |||| | ||| |::||||||||:|:|||||||||| |
Db         58 CTCTGAGTATGCTTACCTAGTCCTCCACGTCCACGCAGCAATATGGAGAGAAAGGGAACT 117

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR---ATTATTAYT 4209
                |||||||||:|||:||||||||||||:||||||||||||||||||:   |||||||:|
Db        118 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGGTTATTATTATT 177

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
                |||:||| |||||||||:||||| ||:|||||||||||||:|||||||||| ||||||||
Db        178 GGCTGTATAGAAACCTAAAGAGGAGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 237

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
                |:||||||||||:|||:|||:||||                 |||||||  ||| |
Db        238 GGAAAGGGAAATAGAAGGGAACTGCC----------------AAAAGAGCCACAAAGTG 280

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
                |||||||||||||||||| ||||  ||   ||| |||||| |||||||||| || ||||
Db        281 GGACCCTCCATTAGAAATGTTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGA 340

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTT 4446
                ||||||||||||:|||||||| | ||| ||||||||||||||| |||   | |||||
Db        341 AACCAAGCCCCAGYACTCAGCAGAATAAATAGAATGGGGAACCTCATGAGGACGTAGTTT 400
```

```
Qy       4447 TCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
                |||||||| ||| || ||||||| |||||||| |||||||||||||||||| ||| |||
Db        401 CCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCA 460

Qy       4507 ATGGAA-ATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATC 4565
                ||||||  |||||||||||||||| |||||||||||||||||| ||||||||||||
Db        461 ATGGAACCTTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATC 520

Qy       4566 ARATGGCCAAATCATTATTTACTGGACCAGG 4596
                |:|||||||||||||||||||||||||||||
Db        521 AGATGGCCAAATCATTATTTACTGGACCAGG 551


RESULT 414
US-09-925-065A-926278
; Sequence 926278, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926278
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-926278

  Query Match             5.4%;  Score 404.1;  DB 1;  Length 552;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  476;  Conservative   26;  Mismatches   42;  Indels   27;  Gaps    5;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
                |||||||:|||:| ||||||||||||    |||||:||||| ||:|| ||||||||||||
Db          1 AATAGCCCTCACTCGGGCACTAGAATT---AGAAGGAAAAAAGGTAAGTATATATACAGA 57

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
                ||||:|:|||||||:|||||||| |||| | ||| |::||||||||:|:|||||||||| |
Db         58 CTCTGAGTATGCTTACCTAGTCCTCCACGTCCACGCAGCAATATGGAGAGAAAGGGAACT 117

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR---ATTATTAYT 4209
                |||||||||:|||:||||||||||||:||||||||||||||||||:   |||||||:|
Db        118 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGGTTATTATTATT 177

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
                |||:||| |||||||||:||||| ||:|||||||||||||:|||||||||| ||||||||
Db        178 GGCTGTATAGAAACCTAAAGAGGAGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 237

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
                |:||||||||||:|||:|||:||||                 |||||||  ||| |
Db        238 GGAAAGGGAAATAGAAGGGAACTGCC----------------AAAAGAGCCACAAAGTG 280

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
                |||||||||||||||||| ||||  ||   ||| |||||| |||||||||| || ||||
Db        281 GGACCCTCCATTAGAAATGTTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGA 340

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTT 4446
                |||||||||||||:|||||||| | ||| ||||||||||||||| |||   | |||||
Db        341 AACCAAGCCCCAGYACTCAGCAGAATAAATAGAATGGGGAACCTCATGAGGACGTAGTTT 400
```

```
Qy       4447 TCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
             |||||||| ||| || |||||||| |||||||| |||||||||||||||||| ||| |||
Db        401 CCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCA 460

Qy       4507 ATGGAA-ATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATC 4565
             ||||||  |||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db        461 ATGGAACCTTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATC 520

Qy       4566 ARATGGCCAAATCATTATTTACTGGACCAGG 4596
             |:||||||||||||||||||||||||||||||
Db        521 AGATGGCCAAATCATTATTTACTGGACCAGG 551


RESULT 415
US-10-717-580-45
; Sequence 45, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-45

  Query Match             5.4%;  Score 403;  DB 1;  Length 425;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  397;  Conservative   10;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7176 CAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTT 7235
             |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 CAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCCAACCTCCCCAACAGCACTTAGGTT 60

Qy       7236 TTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGA 7295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 TTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGA 120

Qy       7296 ATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTT 7355
             ||||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTT 180

Qy       7356 AGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 7415
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 240

Qy       7416 AAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATA 7475
             |||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
Db        241 AAAGAAATAGCCAATCATTTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATA 300

Qy       7476 TAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 7535
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 360

Qy       7536 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||||||||||||||||:||||:
Db        361 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 407


RESULT 416
US-10-717-580-46
; Sequence 46, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
;  APPLICANT: BESEME, Frederic
;  APPLICANT:  BLOND, Jean-Luc
;  APPLICANT:  BOUTON, Olivier
;  APPLICANT:  MANDRAND, Bernard
;  APPLICANT:  MALLET, Francois
;  APPLICANT:  PERRON, Herve
;  TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
;  TITLE OF INVENTION:  WITH PREGNANCY DISORDERS
;  FILE REFERENCE: 105045
;  CURRENT APPLICATION NUMBER: US/10/717,580
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US/09/446,024A
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/FR98/01442
;  PRIOR FILING DATE: 1998-07-06
;  PRIOR APPLICATION NUMBER: FR 97/08815
;  PRIOR FILING DATE: 1997-07-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 427
;   TYPE: DNA
;   ORGANISM: Human
US-10-717-580-46

  Query Match             5.4%;  Score 403;  DB 1;  Length 427;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  397;  Conservative   10;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7176 CAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTT 7235
             |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 CAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTT 60

Qy       7236 TTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGA 7295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 TTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGA 120

Qy       7296 ATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTT 7355
             ||||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTT 180

Qy       7356 AGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 7415
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT 240

Qy       7416 AAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATA 7475
             |||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
Db        241 AAAGAAATAGCCAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATA 300

Qy       7476 TAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 7535
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGG 360

Qy       7536 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||||||||||||||||:||||:
Db        361 GAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 407


RESULT 417
US-09-925-065A-948763
; Sequence 948763, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948763
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948763

  Query Match             5.4%;  Score 402.5;  DB 1;  Length 552;
  Best Local Similarity   83.2%;  Pred. No. 0;
  Matches  475;  Conservative   26;  Mismatches   43;  Indels   27;  Gaps    5;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             ||||||:|||:| |||||||||||    |||||:||||| ||:|| ||||||||||||
Db          1 AATAGCCCTCACTCGGGCACTAGAATT---AGAAGGAAAAAAGGTAAGTATATATACAGA 57

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:||||||:||||||:| |||| | ||| |::||||||||:|:|||||||||| |
Db         58 CTCTGAGTATGCTTACCTAGYCCTCCACGTCCACGCAGCAATATGGAGAGAAAGGGAACT 117

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR---ATTATTAYT 4209
             |||||||||:|||:||||||||||||:|||||||||||||||||||:   |||||||:|
Db        118 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGGTTATTATTATT 177

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
             |||:||| |||||||||:||||| ||:|||||||||||:|||||||||| ||||||||
Db        178 GGCTGTATAGAAACCTAAAGAGGAGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 237

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
             |:||||||||||:|||:|||:||||                ||||||||  ||| |
Db        238 GGAAAGGGAAATAGAAGGGAACTGCC----------------AAAAGAGCCACAAAGTG 280

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
             |||||||||||||||||| ||||  ||   ||| |||||| |||||||||| || ||||
Db        281 GGACCCTCCATTAGAAATGTTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGA 340

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTT 4446
             |||||||||||| |||||||| | ||| |||||||||||||||| |||    | |||||
Db        341 AACCAAGCCCCAGCACTCAGCAGAATAAATAGAATGGGGAACCTCATGAGGACGTAGTTT 400

Qy       4447 TCTCCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
              |||||||| ||| || ||||||| |||||||| |||||||||||||||||| ||| |||
Db        401 CCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCA 460

Qy       4507 ATGGAA-ATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATC 4565
             ||||||  |||||||||||||||| |||||||||||||||||| ||||||||||||
Db        461 ATGGAACCTTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATC 520

Qy       4566 ARATGGCCAAATCATTATTTACTGGACCAGG 4596
             |:|||||||||||||||||||||||||||||
Db        521 AGATGGCCAAATCATTATTTACTGGACCAGG 551


RESULT 418
US-09-925-065A-948763
; Sequence 948763, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948763
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948763

  Query Match             5.4%;  Score 402.5;  DB 1;  Length 552;
  Best Local Similarity   83.2%;  Pred. No. 0;
  Matches  475;  Conservative   26;  Mismatches   43;  Indels   27;  Gaps    5;

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
             ||||||:|||:| |||||||||||    |||||:||||| ||:|| ||||||||||||
Db          1 AATAGCCCTCACTCGGGCACTAGAATT---AGAAGGAAAAAAGGTAAGTATATATACAGA 57

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
             ||||:|:||||||:||||||:| |||| | ||| |::||||||||:|:|||||||||| |
Db         58 CTCTGAGTATGCTTACCTAGYCCTCCACGTCCACGCAGCAATATGGAGAGAAAGGGAACT 117

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR---ATTATTAYT 4209
             |||||||||:|||:||||||||||||:|||||||||||||||||||:   |||||||:|
Db        118 CCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGGTTATTATTATT 177

Qy       4210 GGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAA 4269
             |||:||| |||||||||:||||| ||:|||||||||||:|||||||||| ||||||||
Db        178 GGCTGTATAGAAACCTAAAGAGGAGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAA 237

Qy       4270 GRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCA 4329
             |:||||||||||:|||:|||:||||                ||||||||  ||| |
Db        238 GGAAAGGGAAATAGAAGGGAACTGCC----------------AAAAGAGCCACAAAGTG 280

Qy       4330 GGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGA 4389
             |||||||||||||||||| ||||  ||   ||| |||||| |||||||||| || ||||
Db        281 GGACCCTCCATTAGAAATGTTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGA 340

Qy       4390 AACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG---GCAGTTT 4446
             |||||||||||| |||||||| | ||| |||||||||||||||| |||    | |||||
Db        341 AACCAAGCCCCAGCACTCAGCAGAATAAATAGAATGGGGAACCTCATGAGGACGTAGTTT 400

Qy       4447 TCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCA 4506
              |||||||| ||| || ||||||| |||||||| |||||||||||||||||| ||| |||
Db        401 CCTCCCCTCAGGATGGCTAGCCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCA 460

Qy       4507 ATGGAA-ATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATC 4565
             ||||||  |||||||||||||||| |||||||||||||||||| ||||||||||||
Db        461 ATGGAACCTTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATAGCACCCATC 520

Qy       4566 ARATGGCCAAATCATTATTTACTGGACCAGG 4596
             |:|||||||||||||||||||||||||||||
Db        521 AGATGGCCAAATCATTATTTACTGGACCAGG 551


RESULT 419
US-09-925-065A-341103/c
; Sequence 341103, Application US/09925065A
; Publication No. US20040181048A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341103
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341103

  Query Match             5.4%;  Score 402;  DB 1;  Length 534;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  476;  Conservative    2;  Mismatches   53;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || |||||||||||||| ||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               |  || ||||||  |||||| : |||||||||||||||||| ||||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCYCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| ||||||||| |||||||||| ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             ||||||||||||||||||||||| |    |||||||| |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||| |||||| |||||||||||||||||||||||||||   | | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db          59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 420
US-09-925-065A-341103/c
; Sequence 341103, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341103
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341103

  Query Match             5.4%;  Score 402;  DB 1;  Length 534;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  476;  Conservative    2;  Mismatches   53;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || |||||||||||||| ||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               |  || ||||||  |||||| : |||||||||||||||||| ||||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCYCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| ||||||||| |||||||||| ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             ||||||||||||||||||||||| |    |||||||| |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||| |||||| |||||||||||||||||||||||||||   | | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db          59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 421
US-09-925-065A-341104/c
; Sequence 341104, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341104
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341104

  Query Match             5.4%;  Score 402;  DB 1;  Length 534;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  476;  Conservative    2;  Mismatches   53;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||||  ||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || ||||||||||||||| |||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               | :|| ||||||  ||||||   |||||||||||||||||| ||||||||||||||||||
Db        354 TTTMCCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||| ||| |||||||||  |||||||||||  ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             ||||||||||||||||||||||| |    ||||||||  |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             ||||||||||| |||||| |||||||||||||||||||||||||||   |  | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db         59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 422
US-09-925-065A-341104/c
; Sequence 341104, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341104
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341104

  Query Match             5.4%;  Score 402;  DB 1;  Length 534;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  476;  Conservative    2;  Mismatches   53;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||||  ||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || ||||||||||||||| |||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               | :|| ||||||  ||||||   |||||||||||||||||| ||||||||||||||||||
Db        354 TTTMCCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||| ||| |||||||||  |||||||||||  ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             ||||||||||||||||||||||| |    ||||||||  |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             ||||||||||| |||||| |||||||||||||||||||||||||||   |  | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db         59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 423
US-10-301-480-473604/c
; Sequence 473604, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 473604
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-473604

  Query Match             5.4%;  Score 401.6;  DB 1;  Length 518;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  441;  Conservative   13;  Mismatches   52;  Indels   12;  Gaps    1;

Qy       6710 CAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCT 6769
             ||||||||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||
Db        518 CAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCT 459

Qy       6770 GTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGA 6829
             ||||||| |||||| |||||||||||||||||||| |||||||| |||||| | |||||||
Db        458 GTTTATTATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGA 399

Qy       6830 AAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGAC 6889
             |||||:|||||||| :|||| || |||||| ||:| | ||:||||||| ||||||||||
Db        398 AAGTTAAAGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGAC 339

Qy       6890 CCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAG 6949
              |||||||||   || ||:||||||||| |||||||||||||||||||| ||||:||| ||
Db        338 TCTGGGGCCTTGCCATCCAATGGATGCCTTGGATTCTCCCCTTCTTAGAACCTYTAGTAG 279

Qy       6950 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCT 7009
             ||||||||||||||||||||||||||||||||||||||:||||||||||| |||||||
Db        278 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCT 219

Qy       7010 CTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGT 7057
             |||||||||||  ||||||:|| |||            |||||||||||| ||||||| |
Db        218 CTTCCAGAATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAAT 159

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||| |||||| ||||||| |||| ||||||| ||||| :||||||| | ||||||||||
Db        158 CCATGACTAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTA 99

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             |||| ||  ||| |||||||||||||||||||| |||||||||| |||||| || |||
Db         98 ATGATATGGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCA 39

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACC 7215
             ||||||||||||||| |||||| | |:|| ||||||||
Db         38 ATTCAGCAGGAAGCATTTAGAGTGATCGTTGGCCAACC 1


RESULT 424
US-10-301-480-1087013/c
; Sequence 1087013, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1087013
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1087013

  Query Match             5.4%;  Score 401.6;  DB 1;  Length 518;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  441;  Conservative   13;  Mismatches   52;  Indels   12;  Gaps    1;

Qy       6710 CAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCT 6769
             ||||||||||||:||||| |||||||||||||:||||||| ||||:|| ||||||| ||
Db        518 CAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCT 459

Qy       6770 GTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGA 6829
             ||||||| |||||| |||||||||||||||||||| |||||||| |||||| | |||||||
Db        458 GTTTATTATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGA 399

Qy       6830 AAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGAC 6889
             |||||:|||||||| :|||| || |||||| ||:| | ||:||||||| ||||||||||
Db        398 AAGTTAAAGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGAC 339

Qy       6890 CCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAG 6949
              |||||||||   || ||:||||||||| |||||||||||||||||||| ||||:||| ||
Db        338 TCTGGGGCCTTGCCATCCAATGGATGCCTTGGATTCTCCCCTTCTTAGAACCTYTAGTAG 279

Qy       6950 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCT 7009
             ||||||||||||||||||||||||||||||||||||||:||||||||||| |||||||
Db        278 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCT 219

Qy       7010 CTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGT 7057
             |||||||||||  ||||||:|| |||            |||||||||||| ||||||| |
Db        218 CTTCCAGAATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAAT 159

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||| |||||| ||||||| |||| ||||||| ||||| :||||||| | ||||||||||
Db        158 CCATGACTAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTA 99

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             |||| ||  ||| |||||||||||||||||||| |||||||||| |||||| || |||
Db         98 ATGATATGGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCA 39

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACC 7215
             ||||||||||||||| |||||| | |:|| ||||||||
Db         38 ATTCAGCAGGAAGCATTTAGAGTGATCGTTGGCCAACC 1


RESULT 425
US-10-301-480-455987/c
; Sequence 455987, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 455987
;   LENGTH: 537
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-455987

  Query Match             5.4%;  Score 401.3;  DB 1;  Length 537;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  480;  Conservative    6;  Mismatches   48;  Indels    7;  Gaps    5;

Qy       1425 TTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTC 1484
             || |||||||| ||| ||:||||||||||||||||:|||| ||: |||: ||||| ||
Db        537 TTACTAACTTGCGTGGTAGAAGGACTAAGGAAAACTAGGAAG-AAGCCTATAAATTATTC 479

Qy       1485 AATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 1544
             |||||||||| | |||||| | ||||| ||||||||||||||||||||||||||||||||
Db        478 AATGATGTCCCCTATAACATA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 420
```

```
Qy       1545 TAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAA 1604
             ||||||||||||||||||||  | |||||||  ||||||||||||  ||||||||||||||
Db        419 TAAGGGAGGCATTGAGGAAGTATACCTCTCTCTCACCTGACTCTATTGAAGGCCAACTAA 360

Qy       1605 TCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATT---AGAAAAAACTTCAAAA 1661
             ||||||||  ||||||||||||||||||||||||||||||   | ||||||||||||||
Db        359 TCTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAAAAAAAAAAACTTCAAAA 300

Qy       1662 GTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTT 1721
             |||  || |||| || ||||||||||||:|||||||||||||||||||:| ||||||
Db        299 GTCCACCTTAGGTCCAGAGCAAAACTTARAAACCCTATTGAACTTGGCAACCTTGGTTTT 240

Qy       1722 TTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCA 1781
             |||||||||||||||||||| ||||| |||  |||||||||  | ||| ||||||||||
Db        239 TTATAATAGAGATCAGGAGGAACAGGCAGAATGGGACAAACGACA-TAAGAAAAAGGCCA 181

Qy       1782 CCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCT 1841
             | |||| ||||||| |||||||||||||||||||| |||||||||| | |||||  ||
Db        180 CTGCTTGAGTCATGGCCCTCAGGCAAGTGGACTTTGGTGGCTCTGGAACACGGAAAGCCT 121

Qy       1842 GGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAA 1901
             ||||||||||||||||||||| ||||||||||||| |||||||| ||||||||||||
Db        120 GGGCAAATTGAATGCCTAATAGGGATTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAA 61

Qy       1902 AAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATC 1961
             ||||||||||| ||||| |||||| |||| | ||||||||||||||| ||||||||||
Db         60 AAGATTGTCCAAATAGAAATAAGCCTCCCCCT-GTCCATGCCCCTTATGTCAAGGGAATC 2

Qy       1962 A 1962
              |
Db          1 A 1
```

RESULT 426
US-10-301-480-1069396/c
; Sequence 1069396, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1069396
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1069396

Query Match 5.4%; Score 401.3; DB 1; Length 537;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 480; Conservative 6; Mismatches 48; Indels 7; Gaps 5;

```
Qy       1425 TTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTC 1484
             || |||||||| ||| ||:|||||||||||||||||:|||| ||: |||: ||||| ||
Db        537 TTACTAACTTGCGTGGTAGAAGGACTAAGGAAAACTAGGAAG-AAGCCTATAAATTATTC 479

Qy       1485 AATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 1544
             |||||||||| | |||||| | ||||| ||||||||||||||||||||||||||||||
Db        478 AATGATGTCCCCTATAACATA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 420

Qy       1545 TAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAA 1604
             ||||||||||||||||||  | |||||||  ||||||||||||  ||||||||||||||
Db        419 TAAGGGAGGCATTGAGGAAGTATACCTCTCTCTCACCTGACTCTATTGAAGGCCAACTAA 360
```

```
Qy       1605 TCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATT---AGAAAAAACTTCAAAA 1661
             ||||||||  ||||||||||||||||||||||||||||||   | ||||||||||||||
Db        359 TCTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAAAAAAAAAAACTTCAAAA 300

Qy       1662 GTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTT 1721
             |||  || |||| || ||||||||||||:|||||||||||||||||||:| ||||||
Db        299 GTCCACCTTAGGTCCAGAGCAAAACTTARAAACCCTATTGAACTTGGCAACCTTGGTTTT 240

Qy       1722 TTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCA 1781
             |||||||||||||||||||| ||||| |||  |||||||||  | ||| ||||||||||
Db        239 TTATAATAGAGATCAGGAGGAACAGGCAGAATGGGACAAACGACA-TAAGAAAAAGGCCA 181

Qy       1782 CCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCT 1841
             | |||| ||||||| |||||||||||||||||||| |||||||||| | |||||  ||
Db        180 CTGCTTGAGTCATGGCCCTCAGGCAAGTGGACTTTGGTGGCTCTGGAACACGGAAAGCCT 121

Qy       1842 GGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAA 1901
             ||||||||||||||||||||| ||||||||||||| |||||||| ||||||||||||
Db        120 GGGCAAATTGAATGCCTAATAGGGATTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAA 61

Qy       1902 AAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATC 1961
             ||||||||||| ||||| |||||| |||| | ||||||||||||||| ||||||||||
Db         60 AAGATTGTCCAAATAGAAATAAGCCTCCCCCT-GTCCATGCCCCTTATGTCAAGGGAATC 2

Qy       1962 A 1962
              |
Db          1 A 1
```

RESULT 427
US-09-864-761-7027/c
; Sequence 7027, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30

```
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7027
;   LENGTH: 504
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL049870.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
US-09-864-761-7027

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  |||||||||||||||||||||: |||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||           |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||| |||||| ||||||| ||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 428
US-09-925-065A-341105/c
; Sequence 341105, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341105
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341105

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 534;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  475;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | |||||||| ||| ||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || :|||||||||||||| |||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCAKCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               |  || ||||||  ||||||   |||||||||||||||||| ||||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| |||||||| |||||||||||  ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             |||||||||||||||||||||||| |    |||||||| |||||||||||| | ||||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||||  |||||| ||||||||||||||||||||||||||||||    | | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||||| |||||||||| | ||||||||||||
Db          59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1


RESULT 429
US-09-925-065A-341105/c
; Sequence 341105, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341105
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341105

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 534;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  475;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||||  ||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || :||||||||||||| |||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCAKCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               |  || ||||||  ||||||   |||||||||||||||||| ||||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| ||||||||| |||||||||||| ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             |||||||||||||||||||||||| |    |||||||| |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||||| |||||| ||||||||||||||||||||||||||||    | | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | ||||||||||||||| ||:|||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db          59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 430
US-09-925-065A-341106/c
; Sequence 341106, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341106
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341106

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 534;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  475;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||||  ||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             |||||||||||||||||||||||| ||||||:|||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTRCAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || |||||||||||||| ||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               |  || ||||||  ||||||   |||||||||||||||||| ||||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| ||||||||| |||||||||||| ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             |||||||||||||||||||||||| |    |||||||| |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||||| |||||| ||||||||||||||||||||||||||||    | | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | ||||||||||||||| ||:|||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||||| |||||||||| | ||||||||||||||
Db          59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 431
US-09-925-065A-341106/c
; Sequence 341106, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 341106
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-341106

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 534;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  475;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             | ||||||||| ||| |||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        534 GGCCCTCAGGGAAGCGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAGTTGAAT 475

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             |||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||||
Db        474 GCCTAATAGGGCTTGCTTCCAGTGTGGTCTRCAAGGACACTTTAAAAAAGATTGTCCAAG 415

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             | ||| ||||  ||||| || |||||||||||||| ||||||||||||||||||| ||||
Db        414 TGGAAATAAGTTGCCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCA 355

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
               | || ||||||  ||||||   |||||||||||||||||| ||||||||||||||||
Db        354 TTTACCTAGGGGATGAAGGTCCCCTGAGTCAGAAGCCACTAATCAGATGATCCAGCAGCA 295

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             |||||||||||||||||||||||| ||| ||||||||| |||||||||||| ||    ||
Db        294 GGACTGAGGGTGCCTGGGGCAAGCACCAGCCCATGCCAACACCCTCACAGATCCTCATGT 235

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAG---GTTGTCTCCTGGACACTGGTGCGGTCTTCTT 2151
             ||||||||||||||||||||||| |    |||||||| |||||||||||| | |||||
Db        234 ATGCTTGACCATTGAGGGCCAGGAGGTTAATTGTCTCCGGGACACTGGTGCAGCCTTCTC 175

Qy       2152 AGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTC 2211
             |||||||||||| |||||| |||||||||||||||||||||||||||    |  | |||
Db        174 AGTCTTACTCTCCTGTCCTGGACAACTGTCCTCCAGATCTGTCACTA---TCTGAGGGGT 118

Qy       2212 CNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCT 2271
             | || ||| | |||||||||||||| ||:|||||||||||||||| || |||||||| ||
Db        117 CCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACT 60

Qy       2272 TTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAG 2330
             ||| |||||||||||||||||||||||| |||||||||||| | ||||||||||||||
Db         59 TTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCATTCCCTTGTTAGGGAG 1

RESULT 432
US-10-301-480-319631
; Sequence 319631, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 319631
;   LENGTH: 526
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-319631

  Query Match             5.3%;  Score 399.3;  DB 1;  Length 526;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  423;  Conservative   11;  Mismatches   33;  Indels    1;  Gaps    1;

Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||||| ||||| |||||||||:|||||||||||||||||
Db         60 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 118

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||||| ||||||
Db        119 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 178

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
               ||||| |||||||||:|||:|||||||||||||| ||||||| || |||| |||||||||
Db        179 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 238

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
               |||||||||| |||||||| |  :||| |||| | ||||||||||||||||||||||||||
Db        239 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 298

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
             |||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        299 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 358

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
             |||| |||||||||||||||||||||||||| ||||||||| |||||||||| ||||||
Db        359 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 418

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 478

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTA 6010
             ||||||| |||||||:||||||||| |||:|||||||| ||||| |||
Db        479 ATGTAAAGGAAGTAAYCTCCCAACTGACCCGGGTACATAGCACCCCTA 526


RESULT 433
US-10-301-480-933040
; Sequence 933040, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 933040
;   LENGTH: 526
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-933040

  Query Match             5.3%;  Score 399.3;  DB 1;  Length 526;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  423;  Conservative   11;  Mismatches   33;  Indels    1;  Gaps    1;

Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||||| ||||| |||||||||:|||||||||||||||||
Db         60 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 118

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||||| ||||||
Db        119 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 178
```

```
Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
Db        179 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 238

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
Db        239 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 298

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
Db        299 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 358

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
Db        359 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 418

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
Db        419 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 478

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTA 6010
Db        479 ATGTAAAGGAAGTAAYCTCCCAACTGACCCGGGTACATAGCACCCCTA 526
```

RESULT 434
US-09-925-065A-237209
; Sequence 237209, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 237209
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-237209

Query Match 5.3%; Score 396.9; DB 1; Length 523;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 420; Conservative 11; Mismatches 32; Indels 1; Gaps 1;

```
Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
Db          60 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 118

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
Db         119 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 178

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
Db         179 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 238

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
Db         239 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 298

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
Db         299 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 358

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
Db         359 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 418

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
Db         419 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 478

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACC 6006
Db         479 ATGTAAAGGAAGTAAYCTCCCAACTGACCCGGGTACATAGCACC 522
```

RESULT 435
US-09-925-065A-237209
; Sequence 237209, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 237209
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-237209

Query Match 5.3%; Score 396.9; DB 1; Length 523;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 420; Conservative 11; Mismatches 32; Indels 1; Gaps 1;

```
Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
Db          60 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 118

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
Db         119 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 178

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
Db         179 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 238

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
Db         239 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 298

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
Db         299 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 358

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
```

```
                ||||  |||||||||||||||||||||||||||  |||||||||  ||||||||||  ||||||
Db        359 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 418

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
                ||||||||||  ||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        419 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 478

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACC 6006
                |||||||  |||||||:|||||||||  |||:||||||||  |||||
Db        479 ATGTAAAGGAAGTAAYCTCCCAACTGACCCGGGTACATAGCACC 522


RESULT 436
US-09-925-065A-406935/c
; Sequence 406935, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406935
;   LENGTH: 508
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-406935

  Query Match             5.2%;  Score 393.2;  DB 1;  Length 508;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  432;  Conservative   13;  Mismatches   51;  Indels   12;  Gaps    1;

Qy       6710 CAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCT 6769
                |||||||||||:|||||  |||||||||||||||:|||||||  ||||:||  |||||||  ||
Db        508 CAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCT 449

Qy       6770 GTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGA 6829
                |||||||  ||||||  ||||||||||||||||||||  ||||||||  ||||||  |  |||||||
Db        448 GTTTATTATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGA 389

Qy       6830 AAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGAC 6889
                |||||:||||||||  :||||  ||  ||||||  ||:|  |  ||:|||||||  ||||||||||
Db        388 AAGTTAAAGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGAC 329

Qy       6890 CCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAG 6949
                |||||||||    ||  ||:|||||||||  ||||||||||||||||||||  ||||:|||  ||
Db        328 TCTGGGGCCTTGCCATCCAATGGATGCCTTGGATTCTCCCCTTCTTAGAACCTYTAGTAG 269

Qy       6950 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCT 7009
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||  |||||||
Db        268 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCT 209

Qy       7010 CTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGT 7057
                ||||||||||  ||||||:||  |||            ||||||||||||  |||||||  |
Db        208 CTTCCAGAATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAAT 149

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
                |||  ||||||  |||||||  ||||  |||||||  |||||  :|||||||  |  ||||||||||
Db        148 CCATGACTAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTA 89

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
                ||||  ||  |||  ||||||||||||||||||||||  ||||||||||  ||||||  ||  |||
Db         88 ATGATATGGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCA 29

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
                |||||||||||||||  ||||||  |  |:|
Db         28 ATTCAGCAGGAAGCATTTAGAGTGATCG 1


RESULT 437
US-09-925-065A-406935/c
; Sequence 406935, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406935
;   LENGTH: 508
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-406935

  Query Match             5.2%;  Score 393.2;  DB 1;  Length 508;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  432;  Conservative   13;  Mismatches   51;  Indels   12;  Gaps    1;

Qy       6710 CAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCT 6769
                |||||||||||:|||||  |||||||||||||||:|||||||  ||||:||  |||||||  ||
Db        508 CAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCT 449

Qy       6770 GTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGA 6829
                |||||||  ||||||  ||||||||||||||||||||  ||||||||  ||||||  |  |||||||
Db        448 GTTTATTATTAGGGCAAGAATGCTGTTATTATGTAAATCAATCTGGAATCATGACTGAGA 389

Qy       6830 AAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGAC 6889
                |||||:||||||||  :||||  ||  ||||||  ||:|  |  ||:|||||||  ||||||||||
Db        388 AAGTTAAAGAAATTTGAGATTGATTACAACATAGAACCGAGGAGCTTCAAAACACTGGAC 329

Qy       6890 CCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAG 6949
                |||||||||    ||  ||:|||||||||  ||||||||||||||||||||  ||||:|||  ||
Db        328 TCTGGGGCCTTGCCATCCAATGGATGCCTTGGATTCTCCCCTTCTTAGAACCTYTAGTAG 269

Qy       6950 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCT 7009
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||  |||||||
Db        268 CTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCT 209

Qy       7010 CTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGT 7057
                ||||||||||  ||||||:||  |||            ||||||||||||  |||||||  |
Db        208 CTTCCAGAATCAGAGCTGTAAAGCTACAAATCATTCTTCAAATGGAGCCCCAGATGCAAT 149

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
                |||  ||||||  |||||||  ||||  |||||||  |||||  :|||||||  |  ||||||||||
Db        148 CCATGACTAAAATCTACCACAGATTCCTGGACTGGCCTTCTAGCCCATGCTCTGATGTTA 89

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
                ||||  ||  |||  ||||||||||||||||||||||  ||||||||||  ||||||  ||  |||
Db         88 ATGATATGGAAGCCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCTCCA 29
```

```
Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
              |||||||||||||||| |||||| | |:|
Db         28 ATTCAGCAGGAAGCATTTAGAGTGATCG 1
```

RESULT 438
US-10-301-480-452763
; Sequence 452763, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 452763
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-452763

Query Match 5.2%; Score 390.9; DB 1; Length 488;
Best Local Similarity 91.2%; Pred. No. 0;
Matches 444; Conservative 2; Mismatches 38; Indels 3; Gaps 3;

```
Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
              |||||||||||||||| |||||||||||||||||| ||| ||||||| ||||||||||||||
Db          1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60

Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
              ||||||||| ||||||| |||:| |||||||||||||||||||||||||| ||||||||||
Db         61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
              || |||| ||| ||| |||||| |||||||||||||||||||| ||||| |||||| ||
Db        121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGG 178

Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
              |||||||||||||||| ||||||| |||||||||||   ||||||||||||||||||||
Db        179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
              |||| ||||:|||| |||||||||||||||||||||||| ||||| |||||| || |
Db        239 TCCACTGCGGYCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
               | |||||||||||||| | |||||||||||||||||||||| |||||||| |  ||
Db        299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
              |||| | || ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
              |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db        419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
              |||||||
Db        479 GCCAGGA 485
```

RESULT 439
US-10-301-480-1066172
; Sequence 1066172, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1066172
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1066172

Query Match 5.2%; Score 390.9; DB 1; Length 488;
Best Local Similarity 91.2%; Pred. No. 0;
Matches 444; Conservative 2; Mismatches 38; Indels 3; Gaps 3;

```
Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
              |||||||||||||||| |||||||||||||||||| ||| ||||||| ||||||||||||||
Db          1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60

Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
              ||||||||| ||||||| |||:| |||||||||||||||||||||||||| ||||||||||
Db         61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
              || |||| ||| ||| |||||| |||||||||||||||||||| ||||| |||||| ||
Db        121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGG 178

Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
              |||||||||||||||| ||||||| |||||||||||   ||||||||||||||||||||
Db        179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
              |||| ||||:|||| |||||||||||||||||||||||| ||||| |||||| || |
Db        239 TCCACTGCGGYCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
               | |||||||||||||| | |||||||||||||||||||||| |||||||| |  ||
Db        299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
              |||| | || ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
              |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db        419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
              |||||||
Db        479 GCCAGGA 485
```

RESULT 440
US-10-301-480-452764
; Sequence 452764, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598

```
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 452764
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-452764

  Query Match             5.2%;  Score 389.3;  DB 1;  Length 488;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  443;  Conservative    2;  Mismatches   39;  Indels    3;  Gaps    3;

Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
                ||||||||||||||| |||||||||||||||| ||| ||||||| ||||||||||||||
Db             1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60

Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
                ||||||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||
Db            61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
                || |||| ||| ||| |||||| |||||||||||||||||||| ||||| |||||| ||
Db           121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGG 178

Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
                |||||||||||||||||| ||||||| ||||||||||    |||:|||||||||||||
Db           179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGYCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
                |||| ||||| ||||| |||||||||||||||||||||||| ||||| |||||| || |
Db           239 TCCACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
                | |||||||||||||| | |||||||||||||||||||||||| |||||||| |  ||
Db           299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
                |||| | || ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db           359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
                |||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||
Db           419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
                |||||||
Db           479 GCCAGGA 485


RESULT 441
US-10-301-480-452765
; Sequence 452765, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 452765
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-452765

  Query Match             5.2%;  Score 389.3;  DB 1;  Length 488;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  443;  Conservative    2;  Mismatches   39;  Indels    3;  Gaps    3;

Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
                ||||||||||||||| |||||||||||||||| ||| ||||||| ||||||||||||||
Db             1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60

Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
                ||||||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||
Db            61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
                || |||| ||| ||| |||||| |||||||||||||||||||| ||||| |||||| ||
Db           121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTCR-GCAAGCGG 178

Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
                |||||||||||||||||| ||||||| ||||||||||    |||||||||||||||||||
Db           179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
                |||| ||||| ||||| |||||||||||||||||||||||| ||||| |||||| || |
Db           239 TCCACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
                | |||||||||||||| | |||||||||||||||||||||||| |||||||| |  ||
Db           299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
                |||| | || ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db           359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
                |||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||
Db           419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
                |||||||
Db           479 GCCAGGA 485


RESULT 442
US-10-301-480-1066173
; Sequence 1066173, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1066173
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1066173

  Query Match             5.2%;  Score 389.3;  DB 1;  Length 488;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  443;  Conservative    2;  Mismatches   39;  Indels    3;  Gaps    3;

Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
                ||||||||||||||| |||||||||||||||| ||| ||||||| ||||||||||||||
Db             1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60
```

```
Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
             ||||||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||
Db         61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
             || |||| ||| ||| |||||| |||||||||||||||||||| ||||| |||||| ||
Db        121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGG 178

Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
             |||||||||||||||||| ||||||| ||||||||||   |||:|||||||||||||||
Db        179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGYCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
             |||| ||||| ||||| |||||||||||||||||||||||| ||||| |||||| || |
Db        239 TCCACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
               | ||||||||||||||| | |||||||||||||||||||||||| |||||||| |  ||
Db        299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
             |||| | || ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
             |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db        419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
             |||||||
Db        479 GCCAGGA 485
```

RESULT 443
US-10-301-480-1066174
; Sequence 1066174, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1066174
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1066174

Query Match 5.2%; Score 389.3; DB 1; Length 488;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 443; Conservative 2; Mismatches 39; Indels 3; Gaps 3;

```
Qy       1633 AGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGA 1691
             |||||||||||||||| |||||||||||||||| ||| ||||||| ||||||||||||||
Db          1 AGCTGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGA 60

Qy       1692 AACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGA 1751
             |||||||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||
Db         61 AACCCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGA 120

Qy       1752 ACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGG 1811
             || |||| ||| ||| |||||| |||||||||||||||||||| |||||: |||||| ||
Db        121 ACGGGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTCR-GCAAGCGG 178
```

```
Qy       1812 ACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCT 1871
             |||||||||||||||||| ||||||| ||||||||||   |||||||||||||||||||
Db        179 ACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCT 238

Qy       1872 TCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCC 1931
             |||| ||||| ||||| |||||||||||||||||||||||| ||||| |||||| || |
Db        239 TCCACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTC 298

Qy       1932 TTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAA 1991
               | ||||||||||||||| | |||||||||||||||||||||||| |||||||| |  ||
Db        299 CTTGTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAA 358

Qy       1992 GGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGG 2051
             |||| | || ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        359 GGTCCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCGG 418

Qy       2052 GGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGG 2111
             |||||||||||| |||||||||||||||||||||||| |||||||||||||||||||||||
Db        419 GGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGG 478

Qy       2112 GCCAGGA 2118
             |||||||
Db        479 GCCAGGA 485
```

RESULT 444
US-09-925-065A-567704
; Sequence 567704, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 567704
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-567704

Query Match 5.2%; Score 389.2; DB 1; Length 434;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 382; Conservative 31; Mismatches 21; Indels 0; Gaps 0;

```
Qy       3905 TTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||||
Db          1 TTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTG 60

Qy       3965 GTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTR 4024
             |||||||||||||||| :|||||||||||||| |||||||||||||  ||:| ||:
Db         61 GTCAGTGATAATGGAATCCTTGAAAGTAATCCCCACACTCCAGGAACTAGCACTCAACTG 120

Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
             |||||||||||||||||:|||:|  |||||||||||||||||| :|||||||:|||||||
Db        121 GCAGAACTAATAGCCCTCACTCAGGCACTAGAATTAGGAGAACGAAAAAGGGTAAATATA 180

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
             ||||||||||:|:|||| :|||||||| |||||||||  |::|||||||:|:||||
Db        181 TATACAGACTCTAAGTATGTCTACCTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAA 240
```

```
Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
             ||||||| |||||||||: ||:||||||||||||:||||||||||||||||||:||:||||
Db        241 AGGGAATGCCTAACTTCCAAGGGAACACCTATCAAACATCAGGAAGCCATTARGAGATTA 300

Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAA 4264
             |||:||||:||||||||||||:|||||||:||||||||||||:|||||||||| |||
Db        301 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAA 360

Qy       4265 GGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCA 4324
             ||||||:|||  ||||||:|||:|||: ||||||| :||||||||||:|||||||| |||
Db        361 GGAAAGGAAAAAGAAATAGAAGGGAACAGCCAAGCGGATATTGAAGCCAAAAGAGCCGCA 420

Qy       4325 AGGCAGGACCCTCC 4338
             ||||||||||||||
Db        421 AGGCAGGACCCTCC 434


RESULT 445
US-09-925-065A-567704
; Sequence 567704, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 567704
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-567704

  Query Match             5.2%;  Score 389.2;  DB 1;  Length 434;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  382;  Conservative   31;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       3905 TTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||
Db          1 TTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTG 60

Qy       3965 GTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTR 4024
             |||||||||||||||| :||||||||||||| |||||||||||||  ||:| ||:
Db         61 GTCAGTGATAATGGAATCCTTGAAAGTAATCCCCACACTCCAGGAACTAGCACTCAACTG 120

Qy       4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATA 4084
             ||||||||||||||:|||:|  |||||||||||||||||| :||||||||:|||||||
Db        121 GCAGAACTAATAGCCCTCACTCAGGCACTAGAATTAGGAGAACGAAAAAGGGTAAATATA 180

Qy       4085 TATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAA 4144
             ||||||||||||:|:|||| :|||||||| |||||||||| |::|||||||||:|:||||
Db        181 TATACAGACTCTAAGTATGTCTACCTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAA 240

Qy       4145 AGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204
             ||||||| |||||||||: ||:||||||||||||:||||||||||||||||||:||:||||
Db        241 AGGGAATGCCTAACTTCCAAGGGAACACCTATCAAACATCAGGAAGCCATTARGAGATTA 300

Qy       4205 TTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAA 4264
             |||:||||:||||||||||||:|||||||:||||||||||||:|||||||||| |||
Db        301 TTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATCAGAAA 360

Qy       4265 GGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCA 4324
             ||||||:|||  ||||||:|||:|||: ||||||| :||||||||||:|||||||| |||
Db        361 GGAAAGGAAAAAGAAATAGAAGGGAACAGCCAAGCGGATATTGAAGCCAAAAGAGCCGCA 420

Qy       4325 AGGCAGGACCCTCC 4338
             ||||||||||||||
Db        421 AGGCAGGACCCTCC 434


RESULT 446
US-09-925-065A-295343/c
; Sequence 295343, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295343
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295343

  Query Match             5.2%;  Score 389.2;  DB 1;  Length 437;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  403;  Conservative    5;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6314 ACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTAC 6373
             |||| ||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        436 ACACGACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTAC 377

Qy       6374 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
             ||| |||||||||||||||| ||||| || |||||||||||| ||||||||| |||||
Db        376 CCTTAGGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAG 317

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
             || ||||||||||||||||||||||||||||||:|||:|||| |||||||||||| |||
Db        316 AACCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATT 257

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
             |||||| |||||||||| ||||||||| |||||||| |||||:||||||||||||||||||
Db        256 TATACAATTATGTCATACCTAAGCCCCACAACAAAACAGTACYCATTCTTCCTTTTGTTA 197

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTC 6613
             |  |||||||||||||||||  ||| ||||||||||||||||||||||||||||||||||||
Db        196 TCAGAGCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTC 137

Qy       6614 AGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCC 6673
             |||||||||||||||||||||||||||||||| || |||||||||| |||| |||||||||
Db        136 AGTTCTACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCC 77

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||||||||||||||| |||||||||| ||||||||:|||| ||||||
Db         76 TGGTCACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAG 17

Qy       6734 CTTTAGACTYGCTAAC 6749
```

```
                 |||||||||:||||||
Db         16 CTTTAGACTTGCTAAC 1


RESULT 447
US-09-925-065A-295343/c
; Sequence 295343, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295343
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-295343

  Query Match             5.2%;  Score 389.2;  DB 1;  Length 437;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  403;  Conservative    5;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6314 ACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTAC 6373
             |||| ||||| |||||||||||||||||||||||||||||||||| ||||||||||||||
Db        436 ACACGACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTAC 377

Qy       6374 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
             ||| |||||||||||||||| ||||| || |||||||||||| ||||||||| |||||
Db        376 CCTTAGGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAG 317

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
             || |||||||||||||||||||||||||||||:|||:|||| |||||||||||| |||
Db        316 AACCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATT 257

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
             |||||| |||||||||| ||||||||| |||||||| |||||:||||||||||||||||||
Db        256 TATACAATTATGTCATACCTAAGCCCCACAACAAAACAGTACYCATTCTTCCTTTTGTTA 197

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTC 6613
             |  ||||||||||||||||||   ||| ||||||||||||||||||||||||||||||||
Db        196 TCAGAGCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTC 137

Qy       6614 AGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCC 6673
             |||||||||||||||||||||||||||||||| || |||||||||| |||| |||||||||
Db        136 AGTTCTACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCC 77

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||||||||||||| |||||||||| |||||||||:|||| ||||||
Db         76 TGGTCACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAG 17

Qy       6734 CTTTAGACTYGCTAAC 6749
             |||||||||:||||||
Db         16 CTTTAGACTTGCTAAC 1


RESULT 448
US-10-029-386-23983
; Sequence 23983, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23983
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC008626.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.82
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 6.00e-31
;   OTHER INFORMATION: EST_HUMAN HIT: BF996761.1, EVALUE 1.00e-115
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
US-10-029-386-23983

  Query Match             5.2%;  Score 387.5;  DB 1;  Length 474;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  403;  Conservative   40;  Mismatches   31;  Indels    3;  Gaps    2;

Qy       3344 TAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCA 3403
             |||||||||:|:|||||||:::||| |||||||||||||  ||  | | |||  ||||||
Db          1 TAAGATGGACACCTGAAGCAAAAGAGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCA 58

Qy       3404 AGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAA 3463
             ||:|||| ||||||||:|||||||||:|||||||  ||| |:||:|||::|||||| ||||||
Db         59 AGCCCCAATGTTAAGTTTGCCAACAGGGCAAGGCTTCTCTTTATATGTCACAG-AAAAAA 117

Qy       3464 CAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRY 3523
             |||:||:|||||||:||| |||||||||||:| |:|||||: |||||||||||::||||::
Db        118 CAGGAATAGCTCTAGGAATCCTTACACAGGTTCAAGGGACCAGCTTGCAACCTGTGGCAT 177

Qy       3524 ACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGG 3583
             || |||:|||| |||:|||||||||||||||||||:|:|||||||||: |||||||
Db        178 ACTTGAGTAAGAAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGGTAGTGA 237

Qy       3584 TGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTG 3643
              ||||||||||||::|||||||||||||||||||||||||||||:|||||||||||||
Db        238 CGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTG 297

Qy       3644 TGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAG 3703
             |||||||||||||||:|| ||::| |||||||||||||||||| ||||||||||||||||
Db        298 TGTGGACATCTCATGATGTAAATGGTATACTCACTGCTAAAGGAAACTTGTGGCTGTCAG 357

Qy       3704 ACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCA 3763
             |||||: ||| |||||| :|||||||||||||| ||||||:| | ||||||| :|||| |||
Db        358 ACAACCATTTGCTTAAAAATCAGGCTCTATTAATTGAAGGACTAGTGCTGAGACTGCGCA 417

Qy       3764 CTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAY 3820
             |||||||||||||||||||:| |||||||||||||||||||||||||||| |:||:
Db        418 CTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGAGAGAAC 474


RESULT 449
US-09-925-065A-919227
; Sequence 919227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919227
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919227

  Query Match             5.2%;  Score 386.4;  DB 1;  Length 495;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  423;  Conservative   49;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3014 YGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
                :|||::|||||||||||||||||||||||||||||||||:|||||:|||||||| ||:||
Db          1 TGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTTGCCAC 60

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTT 3133
                ||||||||||:||||||||||:| |:|||||:|||||||| | |||| |||||||||
Db         61 CTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTTT 119

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACT 3193
                |||:||||:|||||||||||||||:|||||||||||||:::|||||||| |||||
Db        120 AGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACT 179

Qy       3194 GGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
                |||||||||||||:|||||||||||||||||||||:|::||| |||||:|| :|||:||
Db        180 GGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTT 239

Qy       3254 CTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTA 3313
                ||||||||:|||||||| |||||||: |||:|||||||| |||:|||||:|||||:|||||
Db        240 CTGCCGAATATGGATT-CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTT 298

Qy       3314 AGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTT 3373
                |||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| :||||
Db        299 AGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCRGCTT 358

Qy       3374 TCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGG 3430
                |||||||||  ||  | | |||  ||||||||:||||||||   ||||:||||||||:||
Db        359 TCCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGG 416

Qy       3431 GCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACAC 3490
                |||||||||| :||:|||::|||||| |||||||||:||:|||||||:|||||||||||||
Db        417 GCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACAC 475

Qy       3491 AGRTCCRAGGGAYGAGCTTG 3510
                ||:|||:|||||: ||| ||
Db        476 AGGTCCAAGGGACAAGCCTG 495


RESULT 450
US-09-925-065A-919227
; Sequence 919227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 919227
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-919227

  Query Match             5.2%;  Score 386.4;  DB 1;  Length 495;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  423;  Conservative   49;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3014 YGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
                :|||::|||||||||||||||||||||||||||||||||:|||||:|||||||| ||:||
Db          1 TGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTTGCCAC 60

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTT 3133
                ||||||||||:||||||||||:| |:|||||:|||||||| | |||| |||||||||
Db         61 CTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTTT 119

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACT 3193
                |||:||||:|||||||||||||||:|||||||||||||:::|||||||| |||||
Db        120 AGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACT 179

Qy       3194 GGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
                |||||||||||||:|||||||||||||||||||||:|::||| |||||:|| :|||:||
Db        180 GGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTT 239

Qy       3254 CTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTA 3313
                ||||||||:|||||||| |||||||: |||:|||||||| |||:|||||:|||||:|||||
Db        240 CTGCCGAATATGGATT-CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTT 298

Qy       3314 AGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTT 3373
                |||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| :||||
Db        299 AGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCRGCTT 358

Qy       3374 TCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGG 3430
                |||||||||  ||  | | |||  ||||||||:||||||||   ||||:||||||||:||
Db        359 TCCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGG 416

Qy       3431 GCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACAC 3490
                |||||||||| :||:|||::|||||| |||||||||:||:|||||||:|||||||||||||
Db        417 GCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACAC 475

Qy       3491 AGRTCCRAGGGAYGAGCTTG 3510
                ||:|||:|||||: ||| ||
Db        476 AGGTCCAAGGGACAAGCCTG 495


RESULT 451
US-09-925-065A-945490
; Sequence 945490, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 945490
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-945490

  Query Match             5.2%;  Score 386.4;  DB 1;  Length 495;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  423;  Conservative   49;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3014 YGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
              :|||::|||||||||||||||||||||||||||||||||:|:|||||:|||||||| ||:||
Db          1 TGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAARCACTCTTAAATTTCCTTGCCAC 60

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTT 3133
              ||||||||||:|||||||||||:| |:|||||:|||||||| | |||| |||||||||
Db         61 CTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTT 119

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACT 3193
              |||:||||:|||||||||||||||||:|||||||||||||||||:::|||||||| |||||
Db        120 AGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACT 179

Qy       3194 GGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
              ||||||||||||||:|||||||||||||||||||||:|::||| |||||:|| :|||:||
Db        180 GGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTT 239

Qy       3254 CTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTA 3313
              |||||||:|||||| ||||||: |||:|||||| |||:|||||:|||||:|||||
Db        240 CTGCCGAATATGGATT-CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTT 298

Qy       3314 AGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTT 3373
              |||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| |||||
Db        299 AGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTT 358

Qy       3374 TCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGG 3430
              |||||||||  ||  | | |||  ||||||||:||||||||    ||||:|||||||:||
Db        359 TCCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGG 416

Qy       3431 GCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACAC 3490
              |||||||||| :||:|||::|||||| |||||||||:||:||||||:|||||||||||||
Db        417 GCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACAC 475

Qy       3491 AGRTCCRAGGGAYGAGCTTG 3510
              ||:|||:|||||: ||| ||
Db        476 AGGTCCAAGGGACAAGCCTG 495


RESULT 452
US-09-925-065A-945490
; Sequence 945490, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 945490
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-945490

  Query Match             5.2%;  Score 386.4;  DB 1;  Length 495;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  423;  Conservative   49;  Mismatches   20;  Indels    8;  Gaps    5;

Qy       3014 YGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYAC 3073
              :|||::|||||||||||||||||||||||||||||||||:|:|||||:|||||||| ||:||
Db          1 TGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAARCACTCTTAAATTTCCTTGCCAC 60

Qy       3074 CTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTT 3133
              ||||||||||:|||||||||||:| |:|||||:|||||||| | |||| |||||||||
Db         61 CTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCTCAGAACAGGCTAAATACTT 119

Qy       3134 AGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACT 3193
              |||:||||:|||||||||||||||||:|||||||||||||||||:::|||||||| |||||
Db        120 AGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACT 179

Qy       3194 GGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTT 3253
              ||||||||||||||:|||||||||||||||||||||:|::||| |||||:|| :|||:||
Db        180 GGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTT 239

Qy       3254 CTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTA 3313
              |||||||:|||||| ||||||: |||:|||||| |||:|||||:|||||:|||||
Db        240 CTGCCGAATATGGATT-CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTT 298

Qy       3314 AGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTT 3373
              |||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| |||||
Db        299 AGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTT 358

Qy       3374 TCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGG 3430
              |||||||||  ||  | | |||  ||||||||:||||||||    ||||:|||||||:||
Db        359 TCCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGG 416

Qy       3431 GCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACAC 3490
              |||||||||| :||:|||::|||||| |||||||||:||:||||||:|||||||||||||
Db        417 GCAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACAC 475

Qy       3491 AGRTCCRAGGGAYGAGCTTG 3510
              ||:|||:|||||: ||| ||
Db        476 AGGTCCAAGGGACAAGCCTG 495


RESULT 453
US-10-301-480-452762/c
; Sequence 452762, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 452762
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-452762
```

```
  Query Match             5.2%;  Score 386.3;  DB 1;  Length 485;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  440;  Conservative    2;  Mismatches   39;  Indels    3;  Gaps    3;

Qy       1636 TGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAAC 1694
             |||||||||||| |||||||||||||||||| ||| ||||||| ||||||||||||||||||
Db        485 TGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAAC 426

Qy       1695 CCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACA 1754
             |||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||||||
Db        425 CCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACG 366

Qy       1755 GGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACT 1814
             |||| ||| ||| |||||| |||||||||||||||||||||| ||||| |||||| |||||
Db        365 GGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACT 308

Qy       1815 TTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCC 1874
             |||||||||||||| ||||||| |||||||||||   ||||||||||||||||||||||
Db        307 TTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCC 248

Qy       1875 AGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTC 1934
             | ||||| ||||| |||||||||||||||||||||||||| ||||| |||||| || | |
Db        247 ACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTT 188

Qy       1935 GTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGT 1994
             |||||||||||||| | |||||||||||||||||||||||| |||||||| |  |||||
Db        187 GTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGT 128

Qy       1995 CTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGC 2054
             | |.|| |||||||||||||||||||||||||||||||||||||||||||||||| :||||
Db        127 CCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCRGGGC 68

Qy       2055 AAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCC 2114
             |||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         67 AAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCC 8

Qy       2115 AGGA 2118
             ||||
Db          7 AGGA 4


RESULT 454
US-10-301-480-1066171/c
; Sequence 1066171, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1066171
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1066171

  Query Match             5.2%;  Score 386.3;  DB 1;  Length 485;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  440;  Conservative    2;  Mismatches   39;  Indels    3;  Gaps    3;

Qy       1636 TGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAAC 1694
             |||||||||||| |||||||||||||||||| ||| ||||||| ||||||||||||||||||
Db        485 TGCAGACATTAGAAAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAAC 426

Qy       1695 CCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACA 1754
             |||||| ||||||| |||:| |||||||||||||||||||||||| ||||||||||||||
Db        425 CCTATTTAACTTGGTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACG 366

Qy       1755 GGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACT 1814
             |||| ||| ||| |||||| |||||||||||||||||||||| ||||| |||||| |||||
Db        365 GGACGAACAGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACT 308

Qy       1815 TTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCC 1874
             |||||||||||||| |||||||||||||||||||   ||||||||||||||||||||||
Db        307 TTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCC 248

Qy       1875 AGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTC 1934
             | ||||| ||||| |||||||||||||||||||||||||| ||||| |||||| || | |
Db        247 ACTGCGGCCTACACGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTT 188

Qy       1935 GTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGT 1994
             |||||||||||||| | |||||||||||||||||||||||| |||||||| |  |||||
Db        187 GTCCATGCCCCTTATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGT 128

Qy       1995 CTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGC 2054
             | | || |||||||||||||||||||||||||||||||||||||||||||||||| :||||
Db        127 CCTCTGGGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCRGGGC 68

Qy       2055 AAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCC 2114
             |||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         67 AAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCC 8

Qy       2115 AGGA 2118
             ||||
Db          7 AGGA 4


RESULT 455
US-10-301-480-372204/c
; Sequence 372204, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 372204
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-372204

  Query Match             5.2%;  Score 385.8;  DB 1;  Length 432;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  399;  Conservative    5;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        431 ACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTTA 372

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||| ||||| || |||||||||||| |||||||| ||||||| ||
Db        371 GGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAGAACCT 312

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||| |||||||||||| ||||||||
Db        311 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATTTATAC 252

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```

```
                | |||||||||| |||||||| |||||||| |||||:||||||||||||||||||  ||
Db        251 AATTATGTCATACCTAAGCCCCACAACAAAACAGTACYCATTCTTCCTTTTGTTATCAGA 192

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  ||| ||||||||||||||||||||||||||||||||||||||
Db        191 GCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 132

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||| || |||||||||| |||| ||||||||||||||
Db        131 TACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCCTGGTC 72

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| |||||||||| ||||||||||:|||| ||||||||||||
Db         71 ACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 12

Qy       6739 GACTYGCTAAC 6749
              ||||:||||||
Db         11 GACTTGCTAAC 1


RESULT 456
US-10-301-480-985613/c
; Sequence 985613, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 985613
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-985613

  Query Match             5.2%;  Score 385.8;  DB 1;  Length 432;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  399;  Conservative    5;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||| |||||||||||||||||||||||||||||||||| |||||||||||||||| |
Db        431 ACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTTA 372

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||| ||||| || |||||||||||| ||||||||| ||||||| ||
Db        371 GGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAGAACCT 312

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||| |||||||||||| ||||||||
Db        311 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATTTATAC 252

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||||| |||||||| |||||||| |||||:||||||||||||||||  ||
Db        251 AATTATGTCATACCTAAGCCCCACAACAAAACAGTACYCATTCTTCCTTTTGTTATCAGA 192

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||  ||| ||||||||||||||||||||||||||||||||||||||
Db        191 GCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 132

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||| || |||||||||| |||| ||||||||||||||
Db        131 TACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCCTGGTC 72

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| |||||||||| |||||||||||:|||| ||||||||||||
Db         71 ACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 12

Qy       6739 GACTYGCTAAC 6749
              ||||:||||||
Db         11 GACTTGCTAAC 1


RESULT 457
US-09-925-065A-387464/c
; Sequence 387464, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 387464
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-387464

  Query Match             5.1%;  Score 385.4;  DB 1;  Length 519;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  453;  Conservative    6;  Mismatches   47;  Indels    6;  Gaps    4;

Qy       1425 TTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTC 1484
              || |||||||| ||| ||:|||||||||||||||||:|||| ||: |||: ||||| ||
Db        519 TTACTAACTTGCGTGGTAGAAGGACTAAGGAAAACTAGGAAG-AAGCCTATAAATTATTC 461

Qy       1485 AATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 1544
              |||||||||| | |||||| | ||||| ||||||||||||||||||||||||||||||||
Db        460 AATGATGTCCCCTATAACATA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 402

Qy       1545 TAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAA 1604
              |||||||||||||||||||||  | ||||||| |||||||||||  ||||||||||||||
Db        401 TAAGGGAGGCATTGAGGAAGTATACCTCTCTCTCACCTGACTCTATTGAAGGCCAACTAA 342

Qy       1605 TCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATT---AGAAAAAACTTCAAAA 1661
              ||||||||  |||||||||||||||||||||||||||||||    | ||||||||||||||
Db        341 TCTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAAAAAAAAAAACTTCAAAA 282

Qy       1662 GTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTT 1721
              |||  || |||| || |||||||||||:|||||||||||||||||||||:| ||||||
Db        281 GTCCACCTTAGGTCCAGAGCAAAACTTARAAACCCTATTGAACTTGGCAACCTTGGTTTT 222

Qy       1722 TTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCA 1781
              |||||||||||||||||||||| ||||| ||| |||||||||  | ||| ||||||||||
Db        221 TTATAATAGAGATCAGGAGGAACAGGCAGAATGGGACAAACGACA-TAAGAAAAAGGCCA 163

Qy       1782 CCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCT 1841
              | |||| |||||||||||||||||||||||||||||||| |||||||||| | |||||  ||
Db        162 CTGCTTGAGTCATGGCCCTCAGGCAAGTGGACTTTGGTGGCTCTGGAACACGGAAAGCCT 103

Qy       1842 GGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAA 1901
              |||||||||||||||||||||||| |||||||||||| |||||||| ||||||||||||
Db        102 GGGCAAATTGAATGCCTAATAGGGATTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAA 43

Qy       1902 AAGATTGTCCAAGTAGAAGTAAGCCGCCCCTT 1933
```

```
                 |||||||||||| ||||| |||||| |||| |
Db         42 AAGATTGTCCAAATAGAAATAAGCCTCCCCCT 11
```

RESULT 458
US-09-925-065A-387464/c
; Sequence 387464, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 387464
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-387464

Query Match 5.1%; Score 385.4; DB 1; Length 519;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 453; Conservative 6; Mismatches 47; Indels 6; Gaps 4;

```
Qy       1425 TTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTC 1484
              || |||||||| ||| ||:|||||||||||||||||:|||| ||: |||: ||||| ||
Db         519 TTACTAACTTGCGTGGTAGAAGGACTAAGGAAAACTAGGAAG-AAGCCTATAAATTATTC 461

Qy       1485 AATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 1544
              |||||||||| | |||||| | ||||| ||||||||||||||||||||||||||||||||
Db         460 AATGATGTCCCCTATAACATA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGAC 402

Qy       1545 TAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAA 1604
              |||||||||||||||||||| | |||||||| |||||||||||  |||||||||||||||
Db         401 TAAGGGAGGCATTGAGGAAGTATACCTCTCTCTCACCTGACTCTATTGAAGGCCAACTAA 342

Qy       1605 TCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATT---AGAAAAAACTTCAAAA 1661
              ||||||||  |||||||||||||||||||||||||||||    | ||||||||||||||
Db         341 TCTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAAAAAAAAAAACTTCAAAA 282

Qy       1662 GTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTT 1721
              |||  || |||| || ||||||||||||:|||||||||||||||||||:| ||||||
Db         281 GTCCACCTTAGGTCCAGAGCAAAACTTARAAACCCTATTGAACTTGGCAACCTTGGTTTT 222

Qy       1722 TTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCA 1781
              |||||||||||||||||||| ||||| |||  |||||||||  | ||| ||||||||||
Db         221 TTATAATAGAGATCAGGAGGAACAGGCAGAATGGGACAAACGACA-TAAGAAAAAGGCCA 163

Qy       1782 CCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCT 1841
              | |||| ||||||| |||||||||||||||||||||| |||||||||  | |||||  ||
Db         162 CTGCTTGAGTCATGGCCCTCAGGCAAGTGGACTTTGGTGGCTCTGGAACACGGAAAGCCT 103

Qy       1842 GGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAA 1901
              |||||||||||||||||||||| |||||||||||| |||||||| ||||||||||||
Db         102 GGGCAAATTGAATGCCTAATAGGGATTGCTTCCAGTGCAGTCTACAAAGACACTTTAAAA 43

Qy       1902 AAGATTGTCCAAGTAGAAGTAAGCCGCCCCTT 1933
              |||||||||||| ||||| |||||| |||| |
Db          42 AAGATTGTCCAAATAGAAATAAGCCTCCCCCT 11
```

RESULT 459
US-09-925-065A-383962/c
; Sequence 383962, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383962
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-383962

Query Match 5.1%; Score 384.8; DB 1; Length 472;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 428; Conservative 2; Mismatches 39; Indels 2; Gaps 2;

```
Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              |||||||||||||||| ||| ||||||| |||||||||||||||||||||| ||||||
Db         472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              | |||:| |||||||||||||||||||||||| |||||||||||||| |||| ||| |||
Db         412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              |||||| |||||||||||||||||||||| ||||| |||||| ||||||||||||||||||
Db         353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              || ||||||| |||||||||||   |||||||||||||||||||||| ||||| |||||
Db         294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db         234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || | |||||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db         174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              |||||||||||||||||||||||||||||||||||||||||||||| :||||||||||| ||||
Db         114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCRGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
              |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4
```

RESULT 460
US-09-925-065A-383962/c
; Sequence 383962, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383962
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-383962

  Query Match             5.1%;  Score 384.8;  DB 1;  Length 472;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  428;  Conservative    2;  Mismatches   39;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             ||||||||||||||||| ||| ||||||| |||||||||||||||||||||||| ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             | |||:| |||||||||||||||||||||||| |||||||||||| |||| ||| |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||| |||||||||||||||||||| ||||| |||||| ||||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             || ||||||| |||||||||||   |||||||||||||||||||||| ||||| |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             |||||||||||||||||||||||||||||||||||||||||| :|||||||||||| ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCRGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
             |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 461
US-09-925-065A-383963/c
; Sequence 383963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383963
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-383963

  Query Match             5.1%;  Score 384.8;  DB 1;  Length 472;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  428;  Conservative    2;  Mismatches   39;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||||||||| ||||||| |||||||||||||||||||||||| ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             | |||:| |||||||||||||||||||||||| |||||||||||| |||| ||| |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||| |||||||||||||||||||| ||||| |||||| ||||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             || ||||||| |||||||||||   |||||||||||||||||||||| ||||| |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGYCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
             |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 462
US-09-925-065A-383963/c
; Sequence 383963, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383963
;    LENGTH: 472
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-383963

  Query Match             5.1%;  Score 384.8;  DB 1;  Length 472;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  428;  Conservative    2;  Mismatches   39;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||| ||| ||||||| ||||||||||||||||||||||||| ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             | |||:| |||||||||||||||||||||||| |||||||||||| |||| ||| |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||| |||||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             || ||||||| |||||||||||    |||||||||||||||||||||||| |||||:|||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGYCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             |||||||||||||||||||||||||||||||||||||||  |||||||||||| ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
             |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 463
US-09-925-065A-383964/c
; Sequence 383964, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383964
;    LENGTH: 472
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-383964

  Query Match             5.1%;  Score 383.2;  DB 1;  Length 472;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  427;  Conservative    2;  Mismatches   40;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             |||||||||||||||||| ||| ||||||| ||||||||||||||||||||||||| ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             | |||:| |||||||||||||||||||||||| |||||||||||| |||| ||| |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||| |||||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             || ||||||| |||||||||||    |||:|||||||||||||||||| ||||| |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGYCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || | |||||||||||||||||||||||| |||||||| |  |||||| | || ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             |||||||||||||||||||||||||||||||||||||||  |||||||||||| ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
             |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 464
US-09-925-065A-383964/c
; Sequence 383964, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383964
;    LENGTH: 472
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-383964

  Query Match             5.1%;  Score 383.2;  DB 1;  Length 472;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  427;  Conservative    2;  Mismatches   40;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
```

```
                |||||||||||||||||  ||| |||||||  ||||||||||||||||||||||||  ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
                | |||:| ||||||||||||||||||||||||  ||||||||||||  ||||  |||  |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
                ||||||  |||||||||||||||||||||||  |||||  ||||||  ||||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-GGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
                ||  |||||||  ||||||||||    |||:|||||||||||||||||||  |||||  |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGYCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
                ||||||||||||||||||||||||  |||||  ||||||  ||  |  |  ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
                ||  |  ||||||||||||||||||||||||  ||||||||  |  ||||||  |  ||  ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
                ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||  ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
                ||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 465
US-09-925-065A-383965/c
; Sequence 383965, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383965
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-383965

  Query Match             5.1%;  Score 383.2;  DB 1;  Length 472;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  427;  Conservative    2;  Mismatches   40;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
                ||||||||||||||||  ||| |||||||  ||||||||||||||||||||||||  ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
                | |||:| ||||||||||||||||||||||||  ||||||||||||  ||||  |||  |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
                ||||||  |||||||||||||||||||||||  |||||  ||||||  ||||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-RGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
                ||  |||||||  ||||||||||    |||:|||||||||||||||||||  |||||  |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
                ||||||||||||||||||||||||  |||||  ||||||  ||  |  |  ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
                ||  |  ||||||||||||||||||||||||  ||||||||  |  ||||||  |  ||  ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
                ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||  ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
                ||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4


RESULT 466
US-09-925-065A-383965/c
; Sequence 383965, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383965
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-383965

  Query Match             5.1%;  Score 383.2;  DB 1;  Length 472;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  427;  Conservative    2;  Mismatches   40;  Indels    2;  Gaps    2;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
                ||||||||||||||||  ||| |||||||  ||||||||||||||||||||||||  ||||||
Db        472 AAAAAACTTCAAAAGTCCGCCTTAGGCCCCGAGCAAAACTTAGAAACCCTATTTAACTTG 413

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
                | |||:| ||||||||||||||||||||||||  ||||||||||||  ||||  |||  |||
Db        412 GTAACCTTGGTTTTTTATAATAGAGATCAGGAAGAGCAGGCGGAACGGGACGAACAGGA- 354

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
                ||||||  |||||||||||||||||||||||  |||||  ||||||  ||||||||||||||||||
Db        353 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCCTC-RGCAAGCGGACTTTGGAGGCTCTGG 295

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
                ||  |||||||  ||||||||||    |||:|||||||||||||||||||  |||||  |||||
Db        294 AACAGGGAAAGGCTGGGCAAATCACATGCCTAATAGGGCTTGCTTCCACTGCGGCCTACA 235

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
                ||||||||||||||||||||||||  |||||  ||||||  ||  |  |  ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
                ||  |  ||||||||||||||||||||||||  ||||||||  |  ||||||  |  ||  ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
                ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||  ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
                ||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4
```

```
Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
              ||||||||||||||||||||||||| ||||| |||||| || | | ||||||||||||||
Db        234 CGGACACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCTCCTTGTCCATGCCCCTT 175

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || | |||||||||||||||||||||||| |||||||  |  |||||| | || ||||||
Db        174 ATGTGAAGGGAATCACTGGAAGGCCCACTTCCCCAGGGAATGAAGGTCCTCTGGGTCAGA 115

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
              |||||||||||||||||||||||||||||||||||||||||||||  |||||||||||| ||||
Db        114 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGCGCCAGCCCA 55

Qy       2068 TGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGA 2118
              |||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          54 TGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGA 4
```

RESULT 467
US-08-979-847-201
; Sequence 201, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 201:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-979-847-201

Query Match 5.1%; Score 382.5; DB 1; Length 438;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 350; Conservative 82; Mismatches 5; Indels 1; Gaps 1;

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              |||||||:||:|::|||||||||||||:||||||:|:||||||:::|||||||:|||:||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              :|:||: ||::|||||||||||||||:||:|||||||||||||:||||||:|||||||:
Db         61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||:|:|||:::|||||||:||||||:::|||||||:|
Db        181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCST 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:|||||:||||:||:||||:||||||||||||:::::|||||||||:||::|
Db        241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:||||||||:||:|||||||  |||:|:|:  |:||::|||||||::||||||:||||::
Db        301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||:|||||||:|||||||:||:|||||||||:|:||||:::::||||
Db        360 ADYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 468
US-10-114-104-201
; Sequence 201, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,847
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A

```
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 201:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 438 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 201:
US-10-114-104-201

  Query Match             5.1%;  Score 382.5;  DB 1;  Length 438;
  Best Local Similarity   79.9%;  Pred. No. 0;
  Matches  350;  Conservative   82;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||:|:||||||:::||||||:|||:||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              :|:||: ||::|||||||||||||||:||:|||||||||||||:||||||:||||||||:
Db         61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||:|:|||:::|||||||:||||||:::|||||||:|
Db        181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCST 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:|||||||:||||:||:||||:||||||||||||:::::|||||||||:||::|
Db        241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:||||||||:||:|||||||  |||:|:|:  |:||::|||||||::||||||:||||::
Db        301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::||||||||||||:||||||||:|||||||:||:|||||||||:|:||||:::::||||
Db        360 ADYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 469
US-08-979-847-128
; Sequence 128, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  128:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-128

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  375;  Conservative   45;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| |||||||:  ||:|::|||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||| |||||||:||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 470
US-08-979-847-129
; Sequence 129, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
```

```
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  129:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-129

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db             1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::||||||||||||||||||||||||||||||||:|||||:||||||||
Db            61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:||||||||||||||||||||
Db           121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||||:  |||||||||||||:::|||||||||
Db           181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db           241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| ||||||||  ||:| |||||||| :|||||:|||||:
Db           301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||:|:|||||:::||||
Db           360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db           420 GGCTTTCCAGGCCCTAAA 437


RESULT 471
US-08-979-847-131
; Sequence 131, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLAC
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  131:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-131

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:|||||||||||:|||  ||||||||||||||
Db             1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::||||||||||||||||||||||||||||||:|||||:||||||||
Db            61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:||||||||||||||||||||
Db           121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||||:  |||||||||||||:::|||||||||
```

```
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| |||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 472
US-10-114-104-128
; Sequence 128, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 128:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-10-114-104-128

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  375;  Conservative   45;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:||| |||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||:||||||||||:|:|:|||||:|||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||:  |||||||||||||:::||||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|::||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| |||||||:|||||||||:|||||||||:|:|||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 473
US-10-114-104-129
; Sequence 129, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
```

```
;             FILING DATE: 26-NOV-1997
;          ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 129:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: cDNA
;          SEQUENCE DESCRIPTION: SEQ ID NO: 129:
US-10-114-104-129

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :||||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:|||||||  ||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 474
US-10-114-104-131
; Sequence 131, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;          CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;          COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;          ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 131:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: cDNA
;          SEQUENCE DESCRIPTION: SEQ ID NO: 131:
US-10-114-104-131

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:|||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :||||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:|||||||  ||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

```
RESULT 475
US-09-925-065A-159044/c
; Sequence 159044, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159044
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159044

  Query Match             5.1%;  Score 379.8;  DB 1;  Length 490;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  429;  Conservative    6;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       4882 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 4941
             |||||||||| ||||||||||||||| |||| |||||| ||||  ||||||||||||||||
Db        490 CCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTA 431

Qy       4942 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 5001
             |||||||| ||||| |||| ||||||||||:|||| |||| |||||||| ||||||||
Db        430 TCCCAGGCCTTAGGCATACAATATCACTTACRCTGCACCTGGAGGCCACAATCCTCAGGA 371

Qy       5002 AAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 5061
             || ||||||||||||||||||: |||| | ||||||||||||||| |||||| ||||||
Db        370 AAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAGAAACC 311

Qy       5062 CACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 5121
             ||||  |||||||| :|||||||| ||||||||||  |||||||| |||| |||||||
Db        310 TACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTCCCCAA 251

Qy       5122 AAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCT 5180
             |||||||||||     ||| | | |||||||||||| |||||||| ||||||||||||
Db        250 AAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGCCCTTCCTAACCAATGACCT 191

Qy       5181 TGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 5239
             |||||||||  | ||    ||||||||||||||||||||||||||||||||||||||
Db        190 TGTGCTTGACAGAGAGCTGACCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 131

Qy       5240 ACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCA 5299
             |||||||||||||||||||| ||||||| | ||||||| || ||||| ||| | |||||
Db        130 ACAAGTTCTTAAAACATTACAGGGAACCTGTTCCTGAGAGGAAGGAAAGGAATTGTTCCA 71

Qy       5300 CCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATC 5359
             |||::|||||||||||||||||||||||:|| |||||||||||||||||||| |||
Db          70 CCCTAGTGACATGGTATTAGTCAAGTCCCTTCCCTTTAATTCCCCATCCCTAGATATATC 11

Qy       5360 CTGGGAAGGA 5369
             ||||||||||
Db          10 CTGGGAAGGA 1


RESULT 476
US-09-925-065A-159044/c
; Sequence 159044, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159044
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159044

  Query Match             5.1%;  Score 379.8;  DB 1;  Length 490;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  429;  Conservative    6;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       4882 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 4941
             |||||||||| ||||||||||||||| |||| |||||| ||||  ||||||||||||||||
Db        490 CCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTA 431

Qy       4942 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 5001
             |||||||| ||||| |||| ||||||||||:|||| |||| |||||||| ||||||||
Db        430 TCCCAGGCCTTAGGCATACAATATCACTTACRCTGCACCTGGAGGCCACAATCCTCAGGA 371

Qy       5002 AAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 5061
             || ||||||||||||||||||: |||| | ||||||||||||||| |||||| ||||||
Db        370 AAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAGAAACC 311

Qy       5062 CACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 5121
             ||||  |||||||| :|||||||| ||||||||||  |||||||| |||| |||||||
Db        310 TACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTCCCCAA 251

Qy       5122 AAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCT 5180
             |||||||||||     ||| | | |||||||||||| |||||||| ||||||||||||
Db        250 AAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGCCCTTCCTAACCAATGACCT 191

Qy       5181 TGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 5239
             |||||||||  | ||    ||||||||||||||||||||||||||||||||||||||
Db        190 TGTGCTTGACAGAGAGCTGACCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 131

Qy       5240 ACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCA 5299
             |||||||||||||||||||| ||||||| | ||||||| || ||||| ||| | |||||
Db        130 ACAAGTTCTTAAAACATTACAGGGAACCTGTTCCTGAGAGGAAGGAAAGGAATTGTTCCA 71

Qy       5300 CCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATC 5359
             |||::|||||||||||||||||||||||||:|| |||||||||||||||||||| |||
Db          70 CCCTAGTGACATGGTATTAGTCAAGTCCCTTCCCTTTAATTCCCCATCCCTAGATATATC 11

Qy       5360 CTGGGAAGGA 5369
             ||||||||||
Db          10 CTGGGAAGGA 1


RESULT 477
US-08-979-847-130
; Sequence 130, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
```

```
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  130:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-130

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:|||||:||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:||||||| ||:||||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||| ||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 478
US-10-114-104-130
; Sequence 130, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 130:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 438 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 130:
US-10-114-104-130

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120
```

```
Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||||:  |||||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||||| ||:||||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              ||||||||||||||| ||||||| ||||||||||:|||||||||:|:||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

```
RESULT 479
US-09-864-761-16318
; Sequence 16318, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16318
;  LENGTH: 433
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: MAP TO AL122001.5
;  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
US-09-864-761-16318

  Query Match             5.1%;  Score 378.6;  DB 1;  Length 433;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  377;  Conservative   38;  Mismatches   17;  Indels    2;  Gaps    1;

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||:||||:||||||||||:|||||:|||| :|||||||||||||||||||||||
Db          2 CCCAGGTATGGCGAAATAGCCAGGCCATTATATACGCTAATTAAGGAAACTCAGAAAGCC 61

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:||||||||:|:|| |||:::|||||||||||||||||  |   |
Db         62 AATACCCATTTAGTAAGATGGACACCTAAAGCAGAAGTGGCTTTCCAGGCCCTAACGAAG 121

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:|||||| |||||:||||||||:||||||||||||:||:|||::
Db        122 GCCCT--AACCCAAGCCCCAGTATTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATG 179

Qy       3451 TCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              |||||||||||||||:||:| ||||:|||||||||||||||:|||:||| |:|||||||
Db        180 TCACAGAAAAAAACAGGAATAACTCTAGGAGTCCTTACACAGGTCCAAGGAACGAGCTTG 239

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
              |||||::||||: ||||||:||||||||:|||||||||||||||||||||:|:|||||||
Db        240 CAACCCGTGGCAGACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 299

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
              |||:||||| |||||||||||||| |::|||||||||||||||||||||||||||||| :
Db        300 TTATGGGTAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTTAAAATAATACAGGAA 359

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
              |||||||||||||||||||||||||||||:||||::|||||||||||||||||||||||
Db        360 AGAGATCTTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGAC 419

Qy       3691 TTGTGGCTGTCAGA 3704
              ||||||||||||||
Db        420 CTGTGGCTGTCAGA 433
```

```
RESULT 480
US-09-864-761-13967
; Sequence 13967, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
```

```
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13967
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AE000660.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
US-09-864-761-13967

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
                ||||||| |||:|:||||||||||: |||||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
                |  |||||||| ||||||||:|||||||||||||||||||||||||||||:|||||:: |
Db            61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
                ||||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||
Db           121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
                ||||||||||||:|||||||||||||:|||| ||||||||| ||||||||||||||:| ||
Db           181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
                |||||||||||||||||||||||||||||:||:|||||||  |||||:|||||  ||||||||
Db           241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
                ||||||  |||||||||||  ||:||||:|||||:||||||||||: ||| ||||||||||
Db           301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
                |||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||||||||||

Db           360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
                |||
Db           420 TGA 422


RESULT 481
US-10-301-480-251973/c
; Sequence 251973, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251973
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-251973

  Query Match             5.0%;  Score 372.8;  DB 1;  Length 483;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches  422;  Conservative    6;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       4882 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 4941
                ||||||||| |||||||||||||| |||| |||||| ||||  ||||||||||||||||||
Db            483 CCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTA 424

Qy       4942 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 5001
                |||||||| ||||| |||| |||||||||||:|||| |||| |||||||| ||||||||
Db            423 TCCCAGGCCTTAGGCATACAATATCACTTACRCTGCACCTGGAGGCCACAATCCTCAGGA 364

Qy       5002 AAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 5061
                || |||||||||||||||||||: |||| | |||||||||||||||| |||||| ||||||
Db            363 AAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAGAAACC 304

Qy       5062 CACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 5121
                ||||  |||||||| :|||||||| |||||||||  ||||||||| |||| |||||||
Db            303 TACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTCCCCAA 244

Qy       5122 AAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCT 5180
                |||||||||||      ||| | | |||||||||||||| |||||||| ||||||||||||||
Db            243 AAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGCCCTTCCTAACCAATGACCT 184

Qy       5181 TGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 5239
                ||||||||||  | ||    |||||||||||||||||||||||||||||||||||||||||
Db            183 TGTGCTTGACAGAGAGCTGACCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 124

Qy       5240 ACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCA 5299
                |||||||||||||||||||| ||||||| | ||||||| || ||||| ||| | |||||
Db           123 ACAAGTTCTTAAAACATTACAGGGAACCTGTTCCTGAGAGGAAGGAAAGGAATTGTTCCA 64

Qy       5300 CCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATC 5359
                |||::|||||||||||||||||||||||||:|| |||||||||||||||||||| |||
Db            63 CCCTAGTGACATGGTATTAGTCAAGTCCCTTCCCTTTAATTCCCCATCCCTAGATATATC 4

Qy       5360 CTG 5362
                |||
Db             3 CTG 1


RESULT 482
US-10-301-480-865382/c
```

```
; Sequence 865382, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865382
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865382

  Query Match             5.0%;  Score 372.8;  DB 1;  Length 483;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches  422;  Conservative    6;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       4882 CCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTA 4941
             ||||||||| ||||||||||||||| |||| |||||| ||||  |||||||||||||||||
Db        483 CCCCGAGGCCTACAGAGTGACAATGGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTA 424

Qy       4942 TCCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGG 5001
             |||||||| ||||| |||| |||||||||||:|||| |||| |||||||| ||||||||
Db        423 TCCCAGGCCTTAGGCATACAATATCACTTACRCTGCACCTGGAGGCCACAATCCTCAGGA 364

Qy       5002 AAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACC 5061
             || ||||||||||||||||||: |||| | |||||||||||||||| |||||| ||||||
Db        363 AAAGTCGAGAAAATGAATGAAATGCTCAGACGACATCTAAAAAAGCGAACCCAAGAAACC 304

Qy       5062 CACCTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAA 5121
               ||||  |||||||| :|||||||| |||||||||  |||||||| |||| |||||||
Db        303 TACCTTGCATGGCCTTCTCTGTTGCTTATAGCCTTACTAAGAATCTGAAACTCTCCCCAA 244

Qy       5122 AAAGCAGGACTT-AGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCT 5180
             ||||||||||||     ||| | | |||||||||||| |||||||| ||||||||||||
Db        243 AAAGCAGGACTTGCATACATGCAAGATGCTGTATGGATGGCCCTTCCTAACCAATGACCT 184

Qy       5181 TGTGCTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 5239
             ||||||||||  | ||    |||||||||||||||||||||||||||||||||||||||
Db        183 TGTGCTTGACAGAGAGCTGACCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCA 124

Qy       5240 ACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCA 5299
             |||||||||||||||||||| ||||||| | ||||||| || ||||| ||| | |||||
Db        123 ACAAGTTCTTAAAACATTACAGGGAACCTGTTCCTGAGAGGAAGGAAAGGAATTGTTCCA 64

Qy       5300 CCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATC 5359
             |||::|||||||||||||||||||||||||:|| |||||||||||||||||||| |||
Db          63 CCCTAGTGACATGGTATTAGTCAAGTCCCTTCCCTTTAATTCCCCATCCCTAGATATATC 4

Qy       5360 CTG 5362
             |||
Db           3 CTG 1


RESULT 483
US-09-925-065A-251122/c
; Sequence 251122, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251122
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-251122

  Query Match             4.9%;  Score 367.8;  DB 1;  Length 471;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  412;  Conservative    9;  Mismatches   41;  Indels   10;  Gaps    2;

Qy       7046 CCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCAC 7105
             ||  |||||||||| ||| ||||||||| |  |||||||||||||||||:|||||||
Db        471 CCCTGATGCAGTCCATGACCAAGATCTACTGTGGACCCCTGGACCGGCCTGCTAGCCCAT 412

Qy       7106 GATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTC 7165
             | |||||||||||||||||||||||||  |||||  |||||||||| |||||||||| |
Db        411 GCTCTGATGTTAATGACATCAAAGGCATGCCTCCCAAGGAAATCTCAACTGCACAACCCC 352

Qy       7166 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAG 7225
             |||||| | |||||||||||||||||||||||||||||||: |||||||||||||| ||||
Db        351 TACTACACTCCAATTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCTCCCCCACAG 292

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||| |||||||||||||||| || |||||||||||||||||| ||||||| ||||||
Db        291 CACTTTGGTTTTCCTGTTGAGAGGGAGGACTGAGAGACAGGACTACCTGGATTCCCTAGG 232

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             |  |:|||||||||:|||||||||:|||||||||||||:| |||||||||||||||||||
Db        231 CCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCRCTTCCACCTTTAAACACGGGG 172

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAA 7396
             |||||||||||:||||||   |||||||          ||||||||||||||||||||||
Db        171 CTTGCAACTTAGCTCACACTCAACCAATCAGAGAGTAAAGAGAGCTCACTAAAATGCTAA 112

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||| |||||||||||||||||||||| |||:||||||:||||||||||  ||
Db        111 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGG 52

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACC 7508
             ||||||||||||  |||||||||||||  |  ||| |||| |||||||| |||
Db          51 AGGGACAATGAT-TGGATATAAACCCAGGCATTCAAGCCAGCAACGGCTACC 1


RESULT 484
US-09-925-065A-251122/c
; Sequence 251122, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251122
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-251122

  Query Match             4.9%;  Score 367.8;  DB 1;  Length 471;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  412;  Conservative    9;  Mismatches   41;  Indels   10;  Gaps    2;

Qy       7046 CCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCAC 7105
             ||  |||||||||| ||| |||||||| |  ||||||||||||||||||:|||||||
Db        471 CCCTGATGCAGTCCATGACCAAGATCTACTGTGGACCCCTGGACCGGCCTGCTAGCCCAT 412

Qy       7106 GATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTC 7165
             | ||||||||||||||||||||||||||  |||||  |||||||||| |||||||||| |
Db        411 GCTCTGATGTTAATGACATCAAAGGCATGCCTCCCAAGGAAATCTCAACTGCACAACCCC 352

Qy       7166 TACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAG 7225
             |||||| | |||||||||||||||||||||||||||||: |||||||||||||| ||||
Db        351 TACTACACTCCAATTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCTCCCCCACAG 292

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||| |||||||||||||||| || |||||||||||||||||||| |||||||| ||||||
Db        291 CACTTTGGTTTTCCTGTTGAGAGGGAGGACTGAGAGACAGGACTACCTGGATTCCCTAGG 232

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             |  |:|||||||||:|||||||||:|||||||||||||:| |||||||||||||||||||
Db        231 CCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCRCTTCCACCTTTAAACACGGGG 172

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||:||||||   |||||||          ||||||||||||||||||||
Db        171 CTTGCAACTTAGCTCACACTCAACCAATCAGAGAGTAAAGAGAGCTCACTAAAATGCTAA 112

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             |||||||| |||||||||||||||||||||||||||||| |||:|||||:||||||||||  ||
Db        111 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGG 52

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACC 7508
             ||||||||||||  |||||||||||||| |  ||| |||| |||||||| |||
Db          51 AGGGACAATGAT-TGGATATAAACCCAGGCATTCAAGCCAGCAACGGCTACC 1


RESULT 485
US-10-301-480-467319/c
; Sequence 467319, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 467319
;   LENGTH: 497
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-467319

  Query Match             4.9%;  Score 365.7;  DB 1;  Length 497;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  394;  Conservative   16;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGG 4266
             |:||||:|||||||||||||:||||||||:|||||||||||||:|||||||||| |||||
Db        497 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGG 438

Qy       4267 AAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAG 4326
             ||||:|||||||||||| |||:|||:  |||||| :||||| ||||:|||| |||  ||||
Db        437 AAAGGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAG 378

Qy       4327 GCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCG 4386
             ||||||| |||||||||||||||||||  ||    ||| ||| || |||||||||| ||||
Db        377 GCAGGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCG 318

Qy       4387 GGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAG 4443
             ||||||||||:|||||||||||||||| ||||| || |||||||| |||||||||    ||
Db        317 GGAAACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAG 258

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||| |||||||| ||| || |||||||  |||||||| ||||||||  |||||||| |||
Db        257 TTTCCTCCCCTCAGGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAA 198

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA 4563
             ||| || |||||||||||||||||||| ||||| |||||||||||||| ||||||||||
Db        197 CCAGTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCA 138

Qy       4564 TCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKT 4623
             |||:|||||||||||||||||| ||||||||||||||||||||||||||||:|||:|
Db        137 TCAGATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGT 78

Qy       4624 CAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||||| ||||:||||||| |:||| || |||||||
Db          77 CAGGGCCTATGAAGTGTGCCACAGAAAGAACCCCCTGC 40


RESULT 486
US-10-301-480-1080728/c
; Sequence 1080728, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1080728
;   LENGTH: 497
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1080728

  Query Match             4.9%;  Score 365.7;  DB 1;  Length 497;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  394;  Conservative   16;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGG 4266
             |:||||:|||||||||||||:||||||||:|||||||||||||:|||||||||| |||||
Db        497 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGG 438

Qy       4267 AAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAG 4326
             ||||:|||||||||||| |||:|||:  |||||| :||||| ||||:|||| |||  ||||
Db        437 AAAGGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAG 378

Qy       4327 GCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCG 4386
             ||||||| |||||||||||||||||||  ||    ||| ||| || |||||||||| ||||
Db        377 GCAGGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCG 318
```

```
Qy       4387 GGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAG 4443
             ||||||||||:|||||||||||||||| ||||| || |||||||| |||||||||     ||
Db        317 GGAAACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAG 258

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||| |||||||| ||| || ||||||| ||||||| |||||||| |||||||| |||
Db        257 TTTCCTCCCCTCAGGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAA 198

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA 4563
             ||| || |||||||||||||||||| ||||| |||||||||||||| ||||||||||
Db        197 CCAGTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCA 138

Qy       4564 TCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKT 4623
             |||:|||||||||||||||||| |||||||||||||||||||||||||||:|||:|
Db        137 TCAGATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGT 78

Qy       4624 CAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||||| ||||:||||||| |:||| || |||||||
Db          77 CAGGGCCTATGAAGTGTGCCACAGAAAGAACCCCCTGC 40


RESULT 487
US-09-925-065A-399976/c
; Sequence 399976, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 399976
;   LENGTH: 500
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-399976

  Query Match             4.9%;  Score 365.7;  DB 1;  Length 500;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  394;  Conservative   16;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGG 4266
             |:||||:|||||||||||||:||||||||:|||||||||||||:||||||||| |||||
Db        500 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGG 441

Qy       4267 AAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAG 4326
             ||||:|||||||||||| |||:|||:  |||||| :||||| ||||:|||| |||  ||||
Db        440 AAAGGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAG 381

Qy       4327 GCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCG 4386
             ||||||| |||||||||||||||||||  ||   ||| ||| || |||||||||| ||||
Db        380 GCAGGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCG 321

Qy       4387 GGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAG 4443
             ||||||||||:|||||||||||||||| ||||| || |||||||| |||||||||     ||
Db        320 GGAAACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAG 261

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||| |||||||| ||| || ||||||| ||||||| |||||||| |||||||| |||
Db        260 TTTCCTCCCCTCAGGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAA 201

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA 4563
             ||| || |||||||||||||||||| ||||| |||||||||||||| ||||||||||
Db        200 CCAGTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCA 141

Qy       4564 TCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKT 4623
             |||:|||||||||||||||||| |||||||||||||||||||||||||||:|||:|
Db        140 TCAGATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGT 81

Qy       4624 CAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGC 4661
             |||||||| ||||:||||||| |:||| || |||||||
Db          80 CAGGGCCTATGAAGTGTGCCACAGAAAGAACCCCCTGC 43


RESULT 488
US-09-925-065A-399976/c
; Sequence 399976, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 399976
;   LENGTH: 500
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-399976

  Query Match             4.9%;  Score 365.7;  DB 1;  Length 500;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  394;  Conservative   16;  Mismatches   45;  Indels    3;  Gaps    1;

Qy       4207 AYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGG 4266
             |:||||:|||||||||||||:||||||||:|||||||||||||:||||||||| |||||
Db        500 ATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGG 441

Qy       4267 AAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAG 4326
             ||||:|||||||||||| |||:|||:  |||||| :||||| ||||:|||| |||  ||||
Db        440 AAAGGAAAGGGAAATAAAAGAGAATCACCAAGCGGATATTAAAGCAAAAACAGCCTCAAG 381

Qy       4327 GCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCG 4386
             ||||||| |||||||||||||||||||  ||   ||| ||| || |||||||||| ||||
Db        380 GCAGGACGCTCCATTAGAAATGCTTATAGAAGGACCCCTAGCATGGGGTAATCCCCTCCG 321

Qy       4387 GGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAG 4443
             ||||||||||:|||||||||||||||| ||||| || |||||||| |||||||||     ||
Db        320 GGAAACCAAGYCCCAGTACTCAGCAGAAGAAATAGGATGGGGAAACTCACGAGGACATAG 261

Qy       4444 TTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTAT 4503
             ||| |||||||| ||| || ||||||| ||||||| |||||||| |||||||| |||
Db        260 TTTCCTCCCCTCAGGAAGGCTAGCCACCAAAGAAGGAAAAATACTCCTGCCTGCAGCTAA 201

Qy       4504 CCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA 4563
             ||| || |||||||||||||||||| ||||| |||||||||||||| ||||||||||
Db        200 CCAGTGAAAATTACTTAAAACCCTTCACCAAACTTTTCACTTAGGCATTGATAGCACCCA 141

Qy       4564 TCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKT 4623
```

```
                ||||:|||||||||||||||||||| ||||||||||||||||||||||||||||||:|||:|
Db        140 TCAGATGGCCAAATCATTATTTACAGGACCAGGCCTTTTCAAAACTATCAAGCAGATAGT 81

Qy       4624 CAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCCTGC 4661
              |||||||| ||||:||||||| |:||| || |||||||
Db         80 CAGGGCCTATGAAGTGTGCCACAGAAAGAACCCCCTGC 43


RESULT 489
US-10-301-480-337623
; Sequence 337623, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 337623
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-337623

  Query Match             4.8%;  Score 361.2;  DB 1;  Length 454;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  405;  Conservative    5;  Mismatches   43;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
              ||||||||||   ||||||:||||||||||||||||:||||||     |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
              ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
              ||||| |||||||||||||||| | |||| ||||||||||||||  ||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATACCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
              |  ||||||  ||||||||||||||||||||||||||||||||||||  |  ||| |
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
              |  || |||||| ||||||||||||||||||||||||||||||:||||||||||
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
              ||||||||||||||||||||:|||| |||||||| ||| ||| ||||||||||||
Db        300 ATAATAGAGATCAGGAGGAGCAGGCRGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
              |||||||||||| |||||||||||||||||||| |||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
              ||||||  ||||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTG 453


RESULT 490
US-10-301-480-951032
; Sequence 951032, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 951032
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-951032

  Query Match             4.8%;  Score 361.2;  DB 1;  Length 454;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  405;  Conservative    5;  Mismatches   43;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
              ||||||||||   ||||||:||||||||||||||||:||||||     |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
              ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
              ||||| |||||||||||||||| | |||| ||||||||||||||  ||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATACCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
              |  ||||||  ||||||||||||||||||||||||||||||||||||  |  ||| |
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
              |  || |||||| ||||||||||||||||||||||||||||||:||||||||||
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
              ||||||||||||||||||||:|||| |||||||| ||| ||| ||||||||||||
Db        300 ATAATAGAGATCAGGAGGAGCAGGCRGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
              |||||||||||| |||||||||||||||||||| |||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
              ||||||  ||||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTG 453


RESULT 491
US-10-301-480-337622
; Sequence 337622, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 337622
```

```
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-337622

  Query Match             4.8%;  Score 359.6;  DB 1;  Length 454;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  404;  Conservative    5;  Mismatches   44;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
              ||||||||||   ||||||:||||||||||||||||:||||||     |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
              ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
              ||||| |||||||||||||| | :||| |||||||||||||  ||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATAMCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
              |  ||||||  ||||||||||||||||||||||||||||||||||||||  |  ||| |
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
              |  || |||||| ||||||||||||||||||||||||||||||||||:||||||||||
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
              |||||||||||||||||||||||| |||| |||||||| ||| ||| ||||||||||||
Db        300 ATAATAGAGATCAGGAGGAGCAGGCAGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
              ||||||||||| |||||||||||||||||||| |||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
              ||||||  |||||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTG 453


RESULT 492
US-10-301-480-951031
; Sequence 951031, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 951031
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-951031

  Query Match             4.8%;  Score 359.6;  DB 1;  Length 454;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  404;  Conservative    5;  Mismatches   44;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
              ||||||||||   ||||||:||||||||||||||||:||||||     |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
              ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
              ||||| |||||||||||||| | :||| |||||||||||||  ||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATAMCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
              |  ||||||  ||||||||||||||||||||||||||||||||||||||  |  ||| |
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
              |  || |||||| ||||||||||||||||||||||||||||||||||:||||||||||
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
              |||||||||||||||||||||||| |||| |||||||| ||| ||| ||||||||||||
Db        300 ATAATAGAGATCAGGAGGAGCAGGCAGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
              ||||||||||| |||||||||||||||||||| |||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTG 1878
              ||||||  |||||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTG 453


RESULT 493
US-09-925-065A-259058
; Sequence 259058, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259058
;   LENGTH: 451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-259058

  Query Match             4.8%;  Score 359.2;  DB 1;  Length 451;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  403;  Conservative    5;  Mismatches   43;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
              ||||||||||   ||||||:||||||||||||||||:||||||     |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
              ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
              ||||| |||||||||||||| | |||| |||||||||||||  ||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATACCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179
```

```
Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
Db        300 ATAATAGAGATCAGGAGGAGCAGGCRGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAG 451
```

RESULT 494
US-09-925-065A-259058
; Sequence 259058, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259058
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-259058

Query Match 4.8%; Score 359.2; DB 1; Length 451;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 403; Conservative 5; Mismatches 43; Indels 2; Gaps 2;

```
Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
Db        120 CTAAGAGAGGCATTGAGGAAGCATACCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
```

```
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
Db        300 ATAATAGAGATCAGGAGGAGCAGGCRGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAG 451
```

RESULT 495
US-09-925-065A-259057
; Sequence 259057, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259057
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-259057

Query Match 4.8%; Score 357.6; DB 1; Length 451;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 402; Conservative 5; Mismatches 44; Indels 2; Gaps 2;

```
Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
Db        120 CTAAGAGAGGCATTGAGGAAGCATAMCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
Db        300 ATAATAGAGATCAGGAGGAGCAGGCAGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
```

```
                 |||||||||||| |||||||||||||||||||| |||||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
                 ||||||  ||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAG 451


RESULT 496
US-09-925-065A-259057
; Sequence 259057, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259057
;   LENGTH: 451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-259057

  Query Match             4.8%;  Score 357.6;  DB 1;  Length 451;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  402;  Conservative    5;  Mismatches   44;  Indels    2;  Gaps    2;

Qy       1424 TTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACT 1483
                 ||||||||||   ||||||:||||||||||||||||:||||||      |: ||||| |
Db          1 TTTGCTAACTGACGTGCTAGAAGGACTAAGGAAAACTAGGAAGAAGCCTAATAAATTAGT 60

Qy       1484 CAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 1543
                 ||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 CAATGATATCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGA 119

Qy       1544 CTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTA 1603
                 ||||| |||||||||||||| | :||| |||||||||||||  |||||||||||||
Db        120 CTAAGAGAGGCATTGAGGAAGCATAMCTCACTGTCACCTGACTCTATTGAAGGCCAACTA 179

Qy       1604 ATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGT 1663
                 |  |||||||  |||||||||||||||||||||||||||||||||||||||  |  ||| |
Db        180 AGGTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAATT 239

Qy       1664 CTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTT 1723
                 |  || |||||| |||||||||||||||||||||||||||||||||||:||||||||||
Db        240 CCACCTTAGGCCTGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTT 299

Qy       1724 ATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACC 1783
                 |||||||||||||||||||||||| |||| |||||||| ||| ||| ||||||||||||
Db        300 ATAATAGAGATCAGGAGGAGCAGGCAGAACGGGACAAACAGGA-TAAGAAAAAGGCCACC 358

Qy       1784 GCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGG 1843
                 |||||||||||| |||||||||||||||||||| |||||||||||| ||||||| ||| |
Db        359 ACTTTAGTCATGGCCCTCAGGCAAGTGGACTTTAGAGGCTCTGGAACAGGGAAAGGCTAG 418

Qy       1844 GCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
                 ||||||  ||||||||||||||||||||||||
Db        419 GCAAATCAAATGCCTAATAGGGCTTGCTTCCAG 451


RESULT 497
US-10-450-763-23371/c
; Sequence 23371, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 23371
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (252)..(335)
;   OTHER INFORMATION: 46% homologous to Escherichia coli microcin J25
;   OTHER INFORMATION: precursor,accession number AF061787,Smith-Waterman Score=60.
US-10-450-763-23371

  Query Match             4.8%;  Score 357.2;  DB 1;  Length 448;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  383;  Conservative    9;  Mismatches   27;  Indels   10;  Gaps    1;

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
                 |||||||||| |||| | |||||| ||||||||||||||| |||||||||:||| | |||
Db         448 CTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCAGTCAA 389

Qy       7214 CCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCT 7273
                 |||||||| |||||||| |||||||||||||||| |||| ||||||||||||||||||||
Db         388 CCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCT 329

Qy       7274 GGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCT 7333
                 |||||||||||||| ||:|||||||||:|||||||||:|||||||||||| | |||||||
Db         328 GGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCT 269

Qy       7334 TTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGC 7393
                 |||||||||||||||||||||||||:||||||| |||||||||||||||||||||||||||
Db         268 TTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGC 209

Qy       7394 TAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGC 7453
                 |||||| ||||| |||||||||||||||||||||| ||||:||| ||:||||||||||||
Db         208 TAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGC 149

Qy       7454 AGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTT 7513
                 | |||||||||||||||||||||||||||||| |  |||||||| |||||||| |||| |||
Db         148 AAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTT 89

Qy       7514 TGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTT 7573
                 |||||||||||||||||||||||| ||||          ||||||||||||||||||||
Db          88 TGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTCACTCTATTAAATCTT 39

Qy       7574 GCARCTGCR 7582
                 |||:||||:
Db          38 GCAACTGCA 30


RESULT 498
US-10-301-480-330802/c
; Sequence 330802, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 330802
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-330802

  Query Match             4.8%;  Score 357;  DB 1;  Length 457;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  400;  Conservative    9;  Mismatches   39;  Indels   10;  Gaps    2;

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             | ||| |||||||||  |  ||||||||||||||||||:|||||||  | ||||||||||||
Db        457 ATGACCAAGATCTACTGTGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTGATGTTAAT 398

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             ||||||||||||||  |||||  |||||||||||| |||||||||| |||||||  | |||||
Db        397 GACATCAAAGGCATGCCTCCCAAGGAAATCTCAACTGCACAACCCCTACTACACTCCAAT 338

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             |||||||||||||||||||||||: ||||||||||||||| ||||||||| ||||||||
Db        337 TCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCTCCCCCACAGCACTTTGGTTTTCC 278

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||| || |||||||||||||||||||| ||||||| |||||||  |:|||||||||
Db        277 TGTTGAGAGGGAGGACTGAGAGACAGGACTACCTGGATTCCCTAGGCCAACTAAGAATCC 218

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||:| |||||||||||||||||||||||||||||||:|
Db        217 CTAAGCCTAGCTGGGAAGGTGACCRCTTCCACCTTTAAACACGGGGCTTGCAACTTAGCT 158

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||   |||||||         |||||||||||||||||||||||||||||||| ||||
Db        157 CACACTCAACCAATCAGAGAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAG 98

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||||||| |||:|||||||:|||||||||||  |||||||||||||||
Db         97 GAGGTAAAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGGAGGGACAATGAT-T 39

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACC 7508
             ||||||||||||| |  ||| |||| |||||||| |||
Db          38 GGATATAAACCCAGGCATTCAAGCCAGCAACGGCTACC 1


RESULT 499
US-10-301-480-944211/c
; Sequence 944211, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944211
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-944211

  Query Match             4.8%;  Score 357;  DB 1;  Length 457;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  400;  Conservative    9;  Mismatches   39;  Indels   10;  Gaps    2;

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             | ||| |||||||||  |  ||||||||||||||||||:|||||||  | ||||||||||||
Db        457 ATGACCAAGATCTACTGTGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTGATGTTAAT 398

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             ||||||||||||||  |||||  |||||||||||| |||||||||| |||||||  | |||||
Db        397 GACATCAAAGGCATGCCTCCCAAGGAAATCTCAACTGCACAACCCCTACTACACTCCAAT 338

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             |||||||||||||||||||||||: ||||||||||||||| ||||||||| ||||||||
Db        337 TCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCTCCCCCACAGCACTTTGGTTTTCC 278

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||| || |||||||||||||||||||| ||||||| |||||||  |:|||||||||
Db        277 TGTTGAGAGGGAGGACTGAGAGACAGGACTACCTGGATTCCCTAGGCCAACTAAGAATCC 218

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||:| |||||||||||||||||||||||||||||||:|
Db        217 CTAAGCCTAGCTGGGAAGGTGACCRCTTCCACCTTTAAACACGGGGCTTGCAACTTAGCT 158

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||   |||||||         |||||||||||||||||||||||||||||||| ||||
Db        157 CACACTCAACCAATCAGAGAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAG 98

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||||||| |||:|||||||:|||||||||||  |||||||||||||||
Db         97 GAGGTAAAGAAATAGCCAACCATCTATTGCCTGAGAGCACAGTGGGAGGGACAATGAT-T 39

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACC 7508
             ||||||||||||| |  ||| |||| |||||||| |||
Db          38 GGATATAAACCCAGGCATTCAAGCCAGCAACGGCTACC 1


RESULT 500
US-09-925-065A-225538/c
; Sequence 225538, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 225538
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-225538

  Query Match             4.7%;  Score 354.5;  DB 1;  Length 440;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  387;  Conservative    7;  Mismatches   37;  Indels    9;  Gaps    1;
```

```
Qy       7119 TGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAA 7178
             ||||||  |||||||||||||| |||||||||||| ||||||  |  | ||||||  |||||||
Db        440 TGACATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGCACCATCCCTACTATGCCCCAA 381

Qy       7179 TTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 7238
             ||||||||||||||||||||||||| |||||||||||||||| |||||||| |||||||
Db        380 TTCAGCAGGAAGCAGTTAGAGCGGTTGTCGGCCAACCTCCCCAGCAGCACTTGGGTTTTC 321

Qy       7239 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATC 7298
             |||||||||  ||||  |||||||||||||||||||||||||  |||||||||:||||||||
Db        320 CTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGATTTTCTAGGCTGACTAAGAATC 261

Qy       7299 CYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGY 7358
             |:|||||||||:||||||||||    | ||||||||||||||||||| ||||||||||:
Db        260 CCTAAGCCTAGCTGGGAAGGTGATGGCTTCCACCTTTAAACACGGGGCCTGCAACTTAGC 201

Qy       7359 TCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACA 7409
             |||||  |  ||||||||         |||||||||||||||||||||||| ||||  |||
Db        200 TCACATCCGACCAATCAGATAGTAAAGAGAGCTCACTAAAATGCTAATTATGCAACAACA 141

Qy       7410 GGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATC 7469
             ||||||||||||||||||||||:||||||:|||||||||||||||||||||||||||||||
Db        140 TGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATC 81

Qy       7470 GGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTT 7529
              ||| || |||||| |  || ||||  |||| :|| |||| |||||||||||||||||||
Db         80 AGGACATGAACCCAGGCATTTGAGCTAGCAAAKGCTACCCTCTTTGGGTCCCCTCCCTTT 21

Qy       7530 GTATGGGAGCTCTGTTTTCA 7549
              ||||||||||||||||||||
Db         20 GTATGGGAGCTCTGTTTTCA 1
```

RESULT 501
US-09-925-065A-225538/c
; Sequence 225538, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 225538
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-225538

Query Match 4.7%; Score 354.5; DB 1; Length 440;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 387; Conservative 7; Mismatches 37; Indels 9; Gaps 1;

```
Qy       7119 TGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAA 7178
             ||||||  |||||||||||||| |||||||||||| ||||||  |  | ||||||  |||||||
Db        440 TGACATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGCACCATCCCTACTATGCCCCAA 381

Qy       7179 TTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTC 7238
             ||||||||||||||||||||||||| |||||||||||||||| |||||||| |||||||
```

```
Db        380 TTCAGCAGGAAGCAGTTAGAGCGGTTGTCGGCCAACCTCCCCAGCAGCACTTGGGTTTTC 321

Qy       7239 CTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATC 7298
             |||||||||  ||||  |||||||||||||||||||||||||  |||||||||:||||||||
Db        320 CTGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGATTTTCTAGGCTGACTAAGAATC 261

Qy       7299 CYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGY 7358
             |:|||||||||:||||||||||    | ||||||||||||||||||| ||||||||||:
Db        260 CCTAAGCCTAGCTGGGAAGGTGATGGCTTCCACCTTTAAACACGGGGCCTGCAACTTAGC 201

Qy       7359 TCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACA 7409
             |||||  |  ||||||||         |||||||||||||||||||||||| ||||  |||
Db        200 TCACATCCGACCAATCAGATAGTAAAGAGAGCTCACTAAAATGCTAATTATGCAACAACA 141

Qy       7410 GGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATC 7469
             ||||||||||||||||||||||:||||||:|||||||||||||||||||||||||||||||
Db        140 TGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATC 81

Qy       7470 GGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTT 7529
              ||| || |||||| |  || ||||  |||| :|| |||| |||||||||||||||||||
Db         80 AGGACATGAACCCAGGCATTTGAGCTAGCAAAKGCTACCCTCTTTGGGTCCCCTCCCTTT 21

Qy       7530 GTATGGGAGCTCTGTTTTCA 7549
              ||||||||||||||||||||
Db         20 GTATGGGAGCTCTGTTTTCA 1
```

RESULT 502
US-09-764-881-183/c
; Sequence 183, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ07
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 183
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-881-183

Query Match 4.7%; Score 353.6; DB 1; Length 439;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 391; Conservative 8; Mismatches 32; Indels 6; Gaps 2;

```
Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||  || ||| ||||||  |||||||||||||||||||||  ||:
Db        439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||| |||||||  ||||||||||||||||  ||||||||||    ||
Db        379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:|||||||||:|||||||||:|||||||| || ||
Db        323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             |||||  |||||||||  ||||||||||||||||:|||||||  |||||||||||||||||
Db        263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||||| ||||||||||||||||||||||||||:|||||:||||
Db        203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| |||||||| ||| |  ||| |||| |||| ||||
Db        143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84
```

```
Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             ||||||||||||||| |||||  |||||||||||||||||||||||  ||||||||||||||
Db         83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db         25 TAAATCTTGCACCTGCA 9


RESULT 503
US-10-242-747-183/c
; Sequence 183, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ07C1
;  CURRENT APPLICATION NUMBER: US/10/242,747
;  CURRENT FILING DATE:  2002-09-13
;  PRIOR APPLICATION NUMBER: 09/764,881
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 183
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-242-747-183

  Query Match             4.7%;  Score 353.6;  DB 1;  Length 439;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   32;  Indels    6;  Gaps    2;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||| || ||| |||||| |||||||||||||||||||||||||| ||:
Db         439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||||| ||||||| ||||||||||||||||  ||||||||||    ||
Db         379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:|||||||||:|||||||||:|||||||||| || ||
Db         323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             ||||| ||||||||| ||||||||||||||||:|||||||  ||||||||||||||||||
Db         263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||| |||||||||||||||||||||||:|||||||:||||
Db         203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| |||||||||| ||| |  ||| |||| |||| ||||
Db         143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             |||||||||||||| |||||  |||||||||||||||||||||||  ||||||||||||||
Db          83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db          25 TAAATCTTGCACCTGCA 9


RESULT 504
US-10-158-057-365/c
; Sequence 365, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ05C1
;  CURRENT APPLICATION NUMBER: US/10/158,057
;  CURRENT FILING DATE:  2002-06-12
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 384
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 365
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-158-057-365

  Query Match             4.7%;  Score 353.6;  DB 1;  Length 439;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  391;  Conservative    8;  Mismatches   32;  Indels    6;  Gaps    2;

Qy       7146 AATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSG 7205
             |||||||||||||||||| || ||| |||||| |||||||||||||||||||||||||| ||:
Db         439 AATCTCAGCTGCACAACCCCTCCTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCAGTCA 380

Qy       7206 TCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAG 7265
             |||||||||||||||||| ||||||| ||||||||||||||||  ||||||||||    ||
Db         379 TCGGCCAACCTCCCCAATAGCACTTGGGTTTTCCTGTTGAGAGTGGGGACTGAG----AG 324

Qy       7266 GACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAC 7325
             |||||||||||||||||||| ||:|||||||||:|||||||||:|||||||||| || ||
Db         323 GACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTAACTAC 264

Qy       7326 ATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCAC 7385
             ||||| ||||||||| ||||||||||||||||:|||||||  ||||||||||||||||||
Db         263 ATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCAC 204

Qy       7386 TAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAG 7445
             |||||||||||||||||| |||||||||||||||||||||||:|||||||:||||
Db         203 TAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAG 144

Qy       7446 AGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCA 7505
             |||||||| |||||||||| |||| |||||||||| ||| |  ||| |||| |||| ||||
Db         143 AGCACAGCGGGAGGGACAAGGATCAGGATATAAATCCAGGCATTCCAGCCAGCAATGGCA 84

Qy       7506 ACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTAT 7565
             |||||||||||||| |||||  |||||||||||||||||||||||  ||||||||||||||
Db          83 ACCCCCTTTGGGTCGCCTCC--TTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTAT 26

Qy       7566 TAAATCTTGCARCTGCR 7582
             |||||||||| ||||:
Db          25 TAAATCTTGCACCTGCA 9


RESULT 505
US-10-301-480-310136
; Sequence 310136, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 310136
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-310136

  Query Match             4.7%;  Score 353.5;  DB 1;  Length 439;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  386;  Conservative    7;  Mismatches   37;  Indels    9;  Gaps    1;

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             |||||  |||||||||||||| |||||||||||| |||||| | | |||||| ||||||||
Db             1 GACATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGCACCATCCCTACTATGCCCCAAT 60

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             |||||||||||||||||||||||| |||||||||||||||| |||||||| ||||||||
Db            61 TCAGCAGGAAGCAGTTAGAGCGGTTGTCGGCCAACCTCCCCAGCAGCACTTGGGTTTTCC 120

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||| |||| |||||||||||||||||||||||| |||||||||:|||||||||
Db           121 TGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGATTTTCTAGGCTGACTAAGAATCC 180

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:||||||||||    | ||||||||||||||||||| ||||||||||:|
Db           181 CTAAGCCTAGCTGGGAAGGTGATGGCTTCCACCTTTAAACACGGGGCCTGCAACTTAGCT 240

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||| | ||||||||         |||||||||||||||||||||| ||||  |||
Db           241 CACATCCGACCAATCAGATAGTAAAGAGAGCTCACTAAAATGCTAATTATGCAACAACAT 300

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||||||||:||||||:||||||||||||||||||||||||||||||
Db           301 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCA 360

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             ||| || |||||| |  || |||| |||| :|| |||| ||||||||||||||||||||
Db           361 GGACATGAACCCAGGCATTTGAGCTAGCAAAKGCTACCCTCTTTGGGTCCCCTCCCTTTG 420

Qy       7531 TATGGGAGCTCTGTTTTCA 7549
             ||||||||||||||||||
Db           421 TATGGGAGCTCTGTTTTCA 439


RESULT 506
US-10-301-480-923545
; Sequence 923545, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923545
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-923545

  Query Match             4.7%;  Score 353.5;  DB 1;  Length 439;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  386;  Conservative    7;  Mismatches   37;  Indels    9;  Gaps    1;

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             |||||  |||||||||||||| |||||||||||| |||||| | | |||||| ||||||||
Db           1 GACATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGCACCATCCCTACTATGCCCCAAT 60

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             |||||||||||||||||||||||| |||||||||||||||| |||||||| ||||||||
Db           61 TCAGCAGGAAGCAGTTAGAGCGGTTGTCGGCCAACCTCCCCAGCAGCACTTGGGTTTTCC 120

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||| |||| |||||||||||||||||||||||| |||||||||:|||||||||
Db           121 TGTTGAGAGGGGGTACTGAGAGACAGGACTAGCTGGATTTTCTAGGCTGACTAAGAATCC 180

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:||||||||||    | ||||||||||||||||||| ||||||||||:|
Db           181 CTAAGCCTAGCTGGGAAGGTGATGGCTTCCACCTTTAAACACGGGGCCTGCAACTTAGCT 240

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||| | ||||||||         |||||||||||||||||||||| ||||  |||
Db           241 CACATCCGACCAATCAGATAGTAAAGAGAGCTCACTAAAATGCTAATTATGCAACAACAT 300

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||||||||:||||||:||||||||||||||||||||||||||||||
Db           301 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCA 360

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             ||| || |||||| |  || |||| |||| :|| |||| ||||||||||||||||||||
Db           361 GGACATGAACCCAGGCATTTGAGCTAGCAAAKGCTACCCTCTTTGGGTCCCCTCCCTTTG 420

Qy       7531 TATGGGAGCTCTGTTTTCA 7549
             ||||||||||||||||||
Db           421 TATGGGAGCTCTGTTTTCA 439


RESULT 507
US-08-979-847-138
; Sequence 138, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  138:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-138

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| || ||||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||        ||||||| |:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:||||||||| |||||||:::|||||||
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||:::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| ||||| |:  ||:|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  |||||||||||||||||||||||||||||| ||:||||||||:|:|||||  :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 508
US-10-114-104-138
; Sequence 138, Application US/10114104
; Publication No. US20030198647A1
;  GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/114,104
;            FILING DATE: 03-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/979,847
;            FILING DATE: 26-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 138:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 429 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-10-114-104-138

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| || ||||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||        ||||||| |:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:||||||||| |||||||:::|||||||
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||:::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| ||||| |:  ||:|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  |||||||||||||||||||||||||||||| ||:||||||||:|:|||||  :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428
```

```
RESULT 509
US-10-301-480-337624
; Sequence 337624, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 337624
;   LENGTH: 481
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-337624

  Query Match             4.6%;  Score 343;  DB 1;  Length 481;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  417;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1923 AGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 1982
             |||| |||| || |||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 AGCCACCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 60

Qy       1983 GGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAG 2042
             ||||||||||||| | |||||||||||||||||||||||| ||||||||||||||||||||
Db         61 GGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATAATCCAGCAGCAGGACTGAG 120

Qy       2043 GGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGA 2102
             |||||||||||||||||||| |||||||||||||||||||||||||  ||||||||||||
Db        121 GGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCAGGTATGCTTGA 180

Qy       2103 CCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTAC 2159
             |||||||| ||||||  | |   | ||| ||||| ||||||| || ||| | ||||||||
Db        181 CCATTGAGTGCCAGGGGGTTAACTCTCTTCTGGATACTGGTGTGGCCTTGTCAGTCTTAC 240

Qy       2160 TCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAC 2219
             ||| ||||||| |||||||||| ||||||||||| |||   |  | |||  | || |||
Db        241 TCTCCTGTCCCAGACAACTGTCTTCCAGATCTGTCTCTA---TCTGAGGGGTCTTAGGAC 297

Qy       2220 GGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTT 2279
             |||||||||||||||| ||:|||||||||||||||| || ||| |||| ||| | ||||
Db        298 AGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTAGGGAACTTCACTCTT 355

Qy       2280 TTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCT 2339
             ||||||||| |||||||||||||| ||||: ||||||| | ||||||||||||||||| ||
Db        356 TTCACATGCCTTTCTAATTATGCCTGAAWTCCCCACTCCTTTGTTAGGGAGAGACATCCT 415

Qy       2340 AGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNC 2399
             ||| |||||||||||||||||||||||||||||||| |||||||| |||||||  ||| | | |
Db        416 AGCTAAAGCAGGGGCCATTATACACCTGAGCATAGGAGAGGGAACACAGGTTATTGTTCC 475

Qy       2400 C 2400
             |
Db        476 C 476


RESULT 510
US-10-301-480-951033
; Sequence 951033, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 951033
;   LENGTH: 481
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-951033

  Query Match             4.6%;  Score 343;  DB 1;  Length 481;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  417;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1923 AGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 1982
             |||| |||| || |||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 AGCCACCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 60

Qy       1983 GGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAG 2042
             ||||||||||||| | |||||||||||||||||||||||| ||||||||||||||||||||
Db         61 GGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATAATCCAGCAGCAGGACTGAG 120

Qy       2043 GGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGA 2102
             |||||||||||||||||||| |||||||||||||||||||||||||  ||||||||||||
Db        121 GGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCAGGTATGCTTGA 180

Qy       2103 CCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTAC 2159
             |||||||| |||||||||||| ||||||||||| |||   |  | |||  | || |||
Db        181 CCATTGAGTGCCAGGGGGTTAACTCTCTTCTGGATACTGGTGTGGCCTTGTCAGTCTTAC 240

Qy       2160 TCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAC 2219
             ||| ||||||| |||||||||| ||||||||||| |||   |  | |||  | || |||
Db        241 TCTCCTGTCCCAGACAACTGTCTTCCAGATCTGTCTCTA---TCTGAGGGGTCTTAGGAC 297

Qy       2220 GGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTT 2279
             |||||||||||||||| ||:|||||||||||||||| || ||| |||| ||| | ||||
Db        298 AGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTAGGGAACTTCACTCTT 355

Qy       2280 TTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCT 2339
             ||||||||| |||||||||||||| ||||: ||||||| | ||||||||||||||||| ||
Db        356 TTCACATGCCTTTCTAATTATGCCTGAAWTCCCCACTCCTTTGTTAGGGAGAGACATCCT 415

Qy       2340 AGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNC 2399
             ||| |||||||||||||||||||||||||||||||| |||||||| |||||||  ||| | | |
Db        416 AGCTAAAGCAGGGGCCATTATACACCTGAGCATAGGAGAGGGAACACAGGTTATTGTTCC 475

Qy       2400 C 2400
             |
Db        476 C 476


RESULT 511
US-09-925-065A-259059
; Sequence 259059, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259059
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-259059

  Query Match             4.6%;  Score 343;  DB 1;  Length 483;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  417;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1923 AGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 1982
             |||| |||| || |||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 AGCCACCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 60

Qy       1983 GGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAG 2042
             ||||||||||||| | ||||||||||||||||||||||| ||||||||||||||||||||
Db         61 GGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATAATCCAGCAGCAGGACTGAG 120

Qy       2043 GGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGA 2102
             |||||||||||||||||||| |||||||||||||||||||||||||  ||||||||||
Db        121 GGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCAGGTATGCTTGA 180

Qy       2103 CCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTAC 2159
             |||||||| ||||||  | |   | ||| ||||| ||||||| || ||| | ||||||||
Db        181 CCATTGAGTGCCAGGGGGTTAACTCTCTTCTGGATACTGGTGTGGCCTTGTCAGTCTTAC 240

Qy       2160 TCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAC 2219
             ||| ||||||| |||||||||| ||||||||||| |||   |  | |||  | || |||
Db        241 TCTCCTGTCCCAGACAACTGTCTTCCAGATCTGTCTCTA---TCTGAGGGGTCTTAGGAC 297

Qy       2220 GGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTT 2279
              |||||||||||||||| ||:|||||||||||||||| || ||| |||| ||| | ||||
Db        298 AGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTAGGGAACTTCACTCTT 355

Qy       2280 TTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCT 2339
             |||||||| |||||||||||| |||| ||||||| | ||||||||||||||||| ||
Db        356 TTCACATGCCTTTCTAATTATGCCTGAAWTCCCCACTCCTTTGTTAGGGAGAGACATCCT 415

Qy       2340 AGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNC 2399
             ||| |||||||||||||||||||||||| |||||||| |||||||  |||  |  | |
Db        416 AGCTAAAGCAGGGGCCATTATACACCTGAGCATAGGAGAGGGAACACAGGTTATTGTTCC 475

Qy       2400 C 2400
             |
Db        476 C 476


RESULT 512
US-09-925-065A-259059
; Sequence 259059, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 259059
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-259059

  Query Match             4.6%;  Score 343;  DB 1;  Length 483;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  417;  Conservative    2;  Mismatches   54;  Indels    8;  Gaps    4;

Qy       1923 AGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 1982
             |||| |||| || |||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 AGCCACCCCCTCATCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCA 60

Qy       1983 GGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAG 2042
             ||||||||||||| | ||||||||||||||||||||||| ||||||||||||||||||||
Db         61 GGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATAATCCAGCAGCAGGACTGAG 120

Qy       2043 GGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGA 2102
             |||||||||||||||||||| |||||||||||||||||||||||||  ||||||||||
Db        121 GGTGCCTGGGGCAAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCAGGTATGCTTGA 180

Qy       2103 CCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTAC 2159
             |||||||| ||||||  | |   | ||| ||||| ||||||| || ||| | ||||||||
Db        181 CCATTGAGTGCCAGGGGGTTAACTCTCTTCTGGATACTGGTGTGGCCTTGTCAGTCTTAC 240

Qy       2160 TCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAC 2219
             ||| ||||||| |||||||||| ||||||||||| |||   |  | |||  | || |||
Db        241 TCTCCTGTCCCAGACAACTGTCTTCCAGATCTGTCTCTA---TCTGAGGGGTCTTAGGAC 297

Qy       2220 GGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTT 2279
              |||||||||||||||| ||:|||||||||||||||| || ||| |||| ||| | ||||
Db        298 AGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTAGGGAACTTCACTCTT 355

Qy       2280 TTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCT 2339
             |||||||| |||||||||||| |||| ||||||| | ||||||||||||||||| ||
Db        356 TTCACATGCCTTTCTAATTATGCCTGAAWTCCCCACTCCTTTGTTAGGGAGAGACATCCT 415

Qy       2340 AGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNC 2399
             ||| |||||||||||||||||||||||| |||||||| |||||||  |||  |  | |
Db        416 AGCTAAAGCAGGGGCCATTATACACCTGAGCATAGGAGAGGGAACACAGGTTATTGTTCC 475

Qy       2400 C 2400
             |
Db        476 C 476


RESULT 513
US-08-979-847-136
; Sequence 136, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
```

```
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22320
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/979,847
;       FILING DATE:  26-NOV-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BERRIDGE, WILLIAM P.
;       REGISTRATION NUMBER:  30,024
;       REFERENCE/DOCKET NUMBER:  WPB 39046A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-836-6400
;       TELEFAX:  703-836-2787
;   INFORMATION FOR SEQ ID NO:  136:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  429 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-979-847-136

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| ||||||||||||||||:|||||:||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:||||| |||||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:||||||||| |||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:|| ||||||||||||||||:|::||||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| ||| | |:  | :||||||||||:||||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 514
US-10-114-104-136
; Sequence 136, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLA
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/114,104
;            FILING DATE: 03-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/979,847
;            FILING DATE: 26-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 136:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 429 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-10-114-104-136

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| ||||||||||||||||:|||||:||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:||||| |||||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:||||||||| |||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:|| ||||||||||||||||:|::||||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| ||| | |:  | :||||||||||:||||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350
```

```
Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              ||||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 515
US-10-301-480-277403/c
; Sequence 277403, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 277403
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-277403

  Query Match             4.5%;  Score 340.7;  DB 1;  Length 453;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  396;  Conservative    7;  Mismatches   40;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| ||||||||||||||||||||| ||
Db        452 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGTGTTAATGACATCAAAGGCATCC 394

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db        393 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 334

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              ||||:||   |||||||||||||||||||||||||| |||||||||||||||| |||||||
Db        333 AGAGYGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 274

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              ||| |||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db        273 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 214

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||||  |||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db        213 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 154

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                      |||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db        153 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 94

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  ||||||||||||||||||||||||||||| |
Db          93 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 35

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCC 7520
              ||| |||||||||||||| |||| ||||||||||
Db          34 ATTCAAGCCGGCAACGGCTACCCTCTTTGGGTCC 1


RESULT 516
US-10-301-480-890812/c
; Sequence 890812, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890812
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-890812

  Query Match             4.5%;  Score 340.7;  DB 1;  Length 453;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  396;  Conservative    7;  Mismatches   40;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| ||||||||||||||||||||| ||
Db        452 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGTGTTAATGACATCAAAGGCATCC 394

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db        393 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 334

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              ||||:||   |||||||||||||||||||||||||| |||||||||||||||| |||||||
Db        333 AGAGYGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 274

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              ||| |||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db        273 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 214

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||||  |||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db        213 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 154

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                      |||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db        153 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 94

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  ||||||||||||||||||||||||||||| |
Db          93 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 35

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCC 7520
              ||| |||||||||||||| |||| ||||||||||
Db          34 ATTCAAGCCGGCAACGGCTACCCTCTTTGGGTCC 1


RESULT 517
US-08-979-847-135
; Sequence 135, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
```

```
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-135

  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches  351;  Conservative   50;  Mismatches   27;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||||:||: ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGAYTTRATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| ||:|||||||||||||::|||||:||||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          |||||||  |:|||||||  |||||||||||  |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAARGCTCAYCTCTGCTCACAYCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||| |:|||:||||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||:||||:|| |||| |||||||||||||:::|||||||||:|| :|
Db        232 ATACTGGRTTATCCYCATCCCAYAACCRTAAAGCAACTAAGARGGTTCCTTGGCATAYCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :|||||||||||:||||||| |||:| |:  |::|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCRGATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::|||| 
Db        351 ARYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 518
US-08-979-847-137
```

```
; Sequence 137, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTI
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  137:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-137

  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  359;  Conservative   39;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| |||||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          |||||||  |:|||||  ||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||  | |||:|||||||||  ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:||||||| ||||||||||||||:|::|||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
```

```
                | :||||||||||:|||||||||||  | |:  | :||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
                | |||||||||||||||||||||||||||||  ||:||||||||:|:||||  :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
                |||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 519
US-10-114-104-135
; Sequence 135, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 135:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-10-114-104-135

  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches  351;  Conservative   50;  Mismatches   27;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
                ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| ||||||:||: ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGAYTTRATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
                |:||: ||::|||||||||||||||| ||:|||||||||||||::|||||:||||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
                ||:|||||||||          ||||||| |:|||||| |||||||||| |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAARGCTCAYCTCTGCTCACAYCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
                ||||||||:||||:|||||||||||| |:|||:|||||||| ||||||:::|||||||:|
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
                ||||||| ||||||:||||:|| |||| |||||||||||||:::|||||||||:|| :|
Db        232 ATACTGGRTTATCCYCATCCCAYAACCRTAAAGCAACTAAGARGGTTCCTTGGCATAYCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
                | :||||||||||:||||||| |||:| |:  |::||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCRGATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
                |::|||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 ARYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
                |||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 520
US-10-114-104-137
; Sequence 137, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYI
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
```

```
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 137:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 429 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-10-114-104-137

  Query Match             4.5%;  Score 339.4;  DB 1;  Length 429;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  359;  Conservative   39;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              |||||||||:||:|::|||||||||||||| |||||| |||||:|||:| |||||| ||| ||
Db             1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
               |:||: ||::|||||||||||||||| |||||||||||||||| :|||||:|||||||||
Db            61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:||||| ||||||||||||||||||||
Db           121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| | |||:|||||||| |||||||:::|||||||| |
Db           172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:||||||| |||||||||||||:|::|||||||||:|||:|
Db           232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||| | |:  | :||||||||||:|||||: |||| |
Db           292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              | |||||||||||||||||||||||||||||| ||:|||||||||:|:||||  :::|||| 
Db           351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db           411 GGCTTTCCAGGCCCTAAA 428


RESULT 521
US-10-301-480-277404/c
; Sequence 277404, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 277404
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-277404

  Query Match             4.5%;  Score 339.1;  DB 1;  Length 453;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  395;  Conservative    7;  Mismatches   41;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| :|||||||||||||||||||| ||
Db           452 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGYGTTAATGACATCAAAGGCATCC 394

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| |||||||||||||||||||||||
Db           393 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 334

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              |||| ||   |||||||||||||||||||||||| |||||||||||||||| |||||||
Db           333 AGAGTGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGAC 274

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              ||| |||||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db           273 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 214

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||||  |||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db           213 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 154

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                      |||||||||||||||||||||||||||||||| ||||||||| ||||||||||
Db           153 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 94

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  |||||||||||||||||||||||||||| |
Db            93 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 35

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCC 7520
              ||| |||||||||||||| |||| ||||||||||
Db            34 ATTCAAGCCGGCAACGGCTACCCTCTTTGGGTCC 1


RESULT 522
US-10-301-480-890813/c
; Sequence 890813, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890813
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-890813

  Query Match             4.5%;  Score 339.1;  DB 1;  Length 453;
  Best Local Similarity   87.0%;  Pred. No. 0;
  Matches  395;  Conservative    7;  Mismatches   41;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| :|||||||||||||||||||| ||
Db           452 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGYGTTAATGACATCAAAGGCATCC 394

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| |||||||||||||||||||||||
Db           393 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 334

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              |||| ||   |||||||||||||||||||||||| |||||||||||||||| |||||||
Db           333 AGAGTGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGAC 274
```

```
Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
             ||| |||||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db        273 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 214

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
             ||||||||  |||||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db        213 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 154

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                     |||||||||||||||||||||||||||||| |||||||||  ||||||||||
Db        153 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 94

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||| | :|||| |||   |  |||||||||||||||||||||||||||||| |
Db         93 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 35

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCC 7520
              ||| |||||||||||||| |||| ||||||||||
Db         34 ATTCAAGCCGGCAACGGCTACCCTCTTTGGGTCC 1


RESULT 523
US-09-880-107-538/c
; Sequence 538, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 538
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA250958
US-09-880-107-538

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             ||||||||||||||  |||||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:|||||||||||| ||||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             |||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||  |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
             ||||||||||||||||||  |||||||||||||||||||||||||:|||
Db          50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 524
US-09-960-706-167/c
; Sequence 167, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 167
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 AA250958
US-09-960-706-167

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             ||||||||||||||  |||||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:|||||||||||| ||||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             |||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||  |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
             ||||||||||||||||||  |||||||||||||||||||||||||:|||
Db          50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 525
US-09-873-319-105/c
; Sequence 105, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
```

```
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 AA250958
US-09-873-319-105

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
             ||||| |||||||||||||||||||| |||: || |||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
             ||||||||||||||  |||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:||| |||||:|||||||||||||| |||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
             |||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
             |||||||||||||||||  |||||||||||||||||||||||||:|||
Db          50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 526
US-09-925-065A-925163
; Sequence 925163, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925163
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925163

  Query Match             4.5%;  Score 338.6;  DB 1;  Length 397;
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  356;  Conservative    2;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              || | |||||||||||||  || |||||||||||||||||||| | ||||||||||||||
Db          9 AAGACACTTCAAAAGTCCACCTTAGGCCCGGAGCAAAACTTAGGAGCCCTATTGAACTTG 68

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              |||||:|| |||||||||||||| |||||||||||||||||| ||||  ||||||||||||
Db         69 GCAACCTCAGTTTTTTATAATAAAGATCAGGAGGAGCAGGTGGAATGGGACAAACGGGAT 128

Qy       1768 TAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              |||||||||||  ||| |||||||||||| |||||||||||| ||||||||||||||||||
Db         129 TAAAAAAAAGGTCACTGCTTTAGTCATGGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGG 188

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         189 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACA 248

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTT 1947
              |||||||||| | |||||||||  || ||||| |||||||||| | ||||||||||||
Db         249 AGGACACTTTTACAAAGATTGTTTAAATAGAAATAAGCCGCCCCCTTGTCCATGCCCCTT 308

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || ||||||||||||||||||||||||||||||||||  || ||: | |||||||||
Db         309 ATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAAGTYCTCTGAGTCAGA 368

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGG 2036
              |||||||||||||||||||||||||||||
Db         369 AGCCACTAACCAGATGATCCAGCAGCAGG 397


RESULT 527
US-09-925-065A-925163
; Sequence 925163, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 925163
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-925163

  Query Match             4.5%;  Score 338.6;  DB 1;  Length 397;
```

```
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  356;  Conservative    2;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
             || | ||||||||||||  || ||||||||||||||||||||| | ||||||||||||||
Db          9 AAGACACTTCAAAAGTCCACCTTAGGCCCGGAGCAAAACTTAGGAGCCCTATTGAACTTG 68

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
             |||||:|| |||||||||||||| |||||||||||||||||| ||||  ||||||||||||
Db         69 GCAACCTCAGTTTTTTATAATAAAGATCAGGAGGAGCAGGTGGAATGGGACAAACGGGAT 128

Qy       1768 TAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
             |||||||||| ||| |||||||||||| |||||||||||| ||||||||||||||||||
Db         129 TAAAAAAAAGGTCACTGCTTTAGTCATGGCCCTCAGGCAAGCGGACTTTGGAGGCTCTGG 188

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         189 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACA 248

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             |||||||||| | |||||||||  || ||||| |||||||||||| | ||||||||||||
Db         249 AGGACACTTTTACAAAGATTGTTTAAATAGAAATAAGCCGCCCCCTTGTCCATGCCCCTT 308

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || ||||||||||||||||||||||||||||||||||  || ||: | |||||||||
Db         309 ATGTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAAGTYCTCTGAGTCAGA 368

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGG 2036
             |||||||||||||||||||||||||||||
Db         369 AGCCACTAACCAGATGATCCAGCAGCAGG 397


RESULT 528
US-09-864-761-3364/c
; Sequence 3364, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3364
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010778.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
US-09-864-761-3364

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             |||||||||||||| | ||||||||:||||||||||:||||||||:||||||||||||
Db          400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db          341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:|||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db          282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db          222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||| ||||||||||||||||:||||||::||||||||||||
Db          162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db          106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||:|||| ||||||||||||||||||||||||:|
Db           46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 529
US-09-925-065A-954031/c
; Sequence 954031, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
```

```
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 954031
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-954031

  Query Match             4.4%;  Score 332.3;  DB 1;  Length 390;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches  359;  Conservative    2;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       1637 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCC 1696
             ||||||||||||||||||||||||||||||  || |||||| ||||||||||||||||||||
Db        390 GCAGACATTAGAAAAAACTTCAAAAGTCCACCTTAGGCCCAGAGCAAAACTTAGAAACCC 331

Qy       1697 TATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGG 1756
             ||||||||||||||||:||||||||||||||||||||||||||||||||||| |||  ||
Db        330 TATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAATGGG 271

Qy       1757 ACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTT 1816
             |||||:||||||  ||||| |||||||||||||||| | ||||||||||   ||||||
Db        270 ACAAAYGGGATT-TAAAAAGGGCCACCGCTTTAGTCACGGCCCTCAGGCAATCAGACTTT 212

Qy       1817 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
             |||||||||||||||||||||||||||||||  ||||||||||||||||  ||||||
Db        211 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTTATTTCCAG 152

Qy       1877 TGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGT 1936
             || |||||||||||||||||||||||||||||||||||||||||| |||||| |||| ||||
Db        151 TGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAATAAGCCACCCCCTCGT 92

Qy       1937 CCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCT 1996
             ||||||||||||| |||||||||||||||||| |||| ||||||||||||  ||||||
Db         91 CCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTGCCCCAGGGGATGAAGGTCC 32

Qy       1997 TTTGAGTCAGAAGCCACTAACCAGATGATCC 2027
             | ||||||||||||||||||||||||||||||
Db         31 TCTGAGTCAGAAGCCACTAACCAGATGATCC 1


RESULT 530
US-09-925-065A-954031/c
; Sequence 954031, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 954031
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-954031

  Query Match             4.4%;  Score 332.3;  DB 1;  Length 390;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches  359;  Conservative    2;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       1637 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCC 1696
             ||||||||||||||||||||||||||||||  || |||||| ||||||||||||||||||||
Db        390 GCAGACATTAGAAAAAACTTCAAAAGTCCACCTTAGGCCCAGAGCAAAACTTAGAAACCC 331

Qy       1697 TATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGG 1756
             ||||||||||||||||:||||||||||||||||||||||||||||||||||| |||  ||
Db        330 TATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAATGGG 271

Qy       1757 ACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTT 1816
             |||||:||||||  ||||| |||||||||||||||| | ||||||||||   ||||||
Db        270 ACAAAYGGGATT-TAAAAAGGGCCACCGCTTTAGTCACGGCCCTCAGGCAATCAGACTTT 212

Qy       1817 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
             |||||||||||||||||||||||||||||||  ||||||||||||||||  ||||||
Db        211 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTTATTTCCAG 152

Qy       1877 TGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGT 1936
             || |||||||||||||||||||||||||||||||||||||||||| |||||| |||| ||||
Db        151 TGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAATAAGCCACCCCCTCGT 92

Qy       1937 CCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCT 1996
             ||||||||||||| |||||||||||||||||| |||| ||||||||||||  ||||||
Db         91 CCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTGCCCCAGGGGATGAAGGTCC 32

Qy       1997 TTTGAGTCAGAAGCCACTAACCAGATGATCC 2027
             | ||||||||||||||||||||||||||||||
Db         31 TCTGAGTCAGAAGCCACTAACCAGATGATCC 1


RESULT 531
US-09-925-065A-948754/c
; Sequence 948754, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948754
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-948754

  Query Match             4.4%;  Score 330.3;  DB 1;  Length 388;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches  357;  Conservative    2;  Mismatches   29;  Indels    1;  Gaps    1;
```

```
Qy       1637 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCC 1696
             ||||||||||||||||||||||||||||  || |||||||| ||||||||||||||||||||
Db        388 GCAGACATTAGAAAAAACTTCAAAAGTCCACCTTAGGCCCAGAGCAAAACTTAGAAACCC 329

Qy       1697 TATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGG 1756
             ||||||||||||||||:||||||||||||||||||||||||||||||||||| |||   ||
Db        328 TATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAATGGG 269

Qy       1757 ACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTT 1816
             |||||:||||||  ||||| |||||||||||||||||| | ||||||||||   ||||||
Db        268 ACAAAYGGGATT-TAAAAAGGGCCACCGCTTTAGTCACGGCCCTCAGGCAATCAGACTTT 210

Qy       1817 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
             |||||||||||||||||||||||||||||||  ||||||||||||||||  ||||||
Db        209 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTTATTTCCAG 150

Qy       1877 TGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGT 1936
             || ||||||||||||||||||||||||||||||||||||||||||| |||||| |||| ||||
Db        149 TGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAATAAGCCACCCCCTCGT 90

Qy       1937 CCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCT 1996
             |||||||||||| |||||||||||||||||| |||| ||||||||||||  ||||||
Db         89 CCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTGCCCCAGGGGATGAAGGTCC 30

Qy       1997 TTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             | ||||||||||||||||||||||||||||
Db          29 TCTGAGTCAGAAGCCACTAACCAGATGAT 1
```

RESULT 532
US-09-925-065A-948754/c
; Sequence 948754, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 948754
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-948754

Query Match 4.4%; Score 330.3; DB 1; Length 388;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 357; Conservative 2; Mismatches 29; Indels 1; Gaps 1;

```
Qy       1637 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCC 1696
             ||||||||||||||||||||||||||||  || |||||||| ||||||||||||||||||||
Db        388 GCAGACATTAGAAAAAACTTCAAAAGTCCACCTTAGGCCCAGAGCAAAACTTAGAAACCC 329

Qy       1697 TATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGG 1756
             ||||||||||||||||:||||||||||||||||||||||||||||||||||| |||   ||
Db        328 TATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAATGGG 269

Qy       1757 ACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTT 1816
             |||||:||||||  ||||| |||||||||||||||||| | ||||||||||   ||||||
Db        268 ACAAAYGGGATT-TAAAAAGGGCCACCGCTTTAGTCACGGCCCTCAGGCAATCAGACTTT 210

Qy       1817 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
             |||||||||||||||||||||||||||||||  ||||||||||||||||  ||||||
Db        209 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTTATTTCCAG 150

Qy       1877 TGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGT 1936
             || ||||||||||||||||||||||||||||||||||||||||||| |||||| |||| ||||
Db        149 TGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAATAAGCCACCCCCTCGT 90

Qy       1937 CCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCT 1996
             |||||||||||| |||||||||||||||||| |||| ||||||||||||  ||||||
Db         89 CCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTGCCCCAGGGGATGAAGGTCC 30

Qy       1997 TTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             | ||||||||||||||||||||||||||||
Db          29 TCTGAGTCAGAAGCCACTAACCAGATGAT 1
```

```
Qy       1817 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAG 1876
             ||||||||||||||||||||||||||||||||  |||||||||||||||  ||||||
Db        209 GGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTTATTTCCAG 150

Qy       1877 TGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGT 1936
             || ||||||||||||||||||||||||||||||||||||||||||| |||||| |||| ||||
Db        149 TGTGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAATAAGCCACCCCCTCGT 90

Qy       1937 CCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCT 1996
             |||||||||||| |||||||||||||||||| |||| ||||||||||||  ||||||
Db         89 CCATGCCCCTTATGTCAAGGGAATCACTGGAAGACCCATTGCCCCAGGGGATGAAGGTCC 30

Qy       1997 TTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             | ||||||||||||||||||||||||||||
Db          29 TCTGAGTCAGAAGCCACTAACCAGATGAT 1
```

RESULT 533
US-09-925-065A-187359/c
; Sequence 187359, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 187359
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-187359

Query Match 4.4%; Score 329.3; DB 1; Length 440;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 384; Conservative 7; Mismatches 39; Indels 11; Gaps 3;

```
Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
             || |||||||||||| |||||  |||||| | |||| |||||||||||||||||||| ||
Db        439 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGTGTTAATGACATCAAAGGCATCC 381

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
             |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db        380 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 321

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
             ||||:||  |||||||||||||||||||||||| |||||||||||||| |||||||
Db        320 AGAGYGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 261

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
             ||| ||||||||||||||||||||||||||||||||||||:|||| ||||:|||||||:|||||
Db        260 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 201

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
             ||||||||  |||||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db        200 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 141

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                     |||||||||||||||||||||||||||||||||||||| |||||||||  ||||||||||
```

```
Db        140 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 81

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  |||||||||||||||||||||||||||||| |
Db         80 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 22

Qy       7487 YTTCGAGCCGGCAACGGCAAC 7507
                 ||| |||||||||||||  ||
Db         21 ATTCAAGCCGGCAACGGCTAC 1


RESULT 534
US-09-925-065A-187359/c
; Sequence 187359, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 187359
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-187359

  Query Match             4.4%;  Score 329.3;  DB 1;  Length 440;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  384;  Conservative    7;  Mismatches   39;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| ||||||||||||||||||||| ||
Db         439 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGTGTTAATGACATCAAAGGCATCC 381

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db         380 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 321

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              ||||:||   |||||||||||||||||||||||| |||||||||||||| |||||||
Db         320 AGAGYGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 261

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              ||| |||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db         260 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 201

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||||  |||||||||||||| |||||||| |||||||:|||||  |  |||||||
Db         200 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 141

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                     |||||||||||||||||||||||||||||||| ||||||||| ||||||||||
Db         140 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 81

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  |||||||||||||||||||||||||||||| |
Db         80 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 22

Qy       7487 YTTCGAGCCGGCAACGGCAAC 7507
                 ||| |||||||||||||  ||
Db         21 ATTCAAGCCGGCAACGGCTAC 1


RESULT 535
US-09-925-065A-870085
; Sequence 870085, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870085
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-870085

  Query Match             4.4%;  Score 328.3;  DB 1;  Length 448;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  379;  Conservative    2;  Mismatches   36;  Indels    3;  Gaps    3;

Qy       2207 GGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGG 2266
              |||  | || ||| | |||||||||||||| ||:|||||||||||||||| || ||||||
Db         12 GGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGG 69

Qy       2267 GAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAG 2326
              || |||||| |||||||||||||||||||||||| |||||||||||| |||||||||
Db         70 GAACTTTATACTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAG 129

Qy       2327 GGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACAC 2386
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        130 GGAGAGACATTCTAGCRAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAAGAACAC 189

Qy       2387 CCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 2446
              |||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CCGTTTATTGTCCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 249

Qy       2447 CAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTC 2506
              ||| ||||||||| |||||||||||| | |||||  |||||||||||||||||| ||| ||
Db        250 CAACATGGACGAG-CAAAGAATGCCCATTCTGTTTGAGTTAAACTAAAGGATTTCACCTC 308

Qy       2507 CTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGAT 2566
              |||||||||||||||||||||||||||||| | |||||||||||||||||| || |||| ||
Db        309 CTTTCCCTACCAAAGGCAGTACCCCCTTACACCCAAGGCCCAACAAGGACTCTAAAATAT 368

Qy       2567 TGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTC 2626
              | |||||||| |||||||||||||| ||||||||||||||||   |  ||||||| || | |
Db        369 TTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC 428


RESULT 536
US-09-925-065A-870085
; Sequence 870085, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
```

```
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870085
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-870085

  Query Match             4.4%;  Score 328.3;  DB 1;  Length 448;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  379;  Conservative    2;  Mismatches   36;  Indels    3;  Gaps    3;

Qy       2207 GGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGG 2266
             |||   | || ||| | |||||||||||||| ||:|||||||||||||||| || ||||||
Db         12 GGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGG 69

Qy       2267 GAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAG 2326
             || |||||| |||||||||||||||||||||||||||| |||||||||||| |||||||||
Db         70 GAACTTTATACTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAG 129

Qy       2327 GGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACAC 2386
             ||||||||||||||:|||||||||||||||||||||||||||||||||||||||| ||||||
Db        130 GGAGAGACATTCTAGCRAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAAGAACAC 189

Qy       2387 CCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 2446
             |||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CCGTTTATTGTCCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 249

Qy       2447 CAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTC 2506
             ||| ||||||||| |||||||||||| | |||||  |||||||||||||||||| ||| ||
Db        250 CAACATGGACGAG-CAAAGAATGCCCATTCTGTTTGAGTTAAACTAAAGGATTTCACCTC 308

Qy       2507 CTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGAT 2566
             |||||||||||||||||||||||||| | |||||||||||||||||| || |||| ||
Db        309 CTTTCCCTACCAAAGGCAGTACCCCCTTACACCCAAGGCCCAACAAGGACTCTAAAATAT 368

Qy       2567 TGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTC 2626
             | |||||||| |||||||||||||| |||||||||||||||  |  ||||||| || | |
Db        369 TTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC 428


RESULT 537
US-09-925-065A-187360/c
; Sequence 187360, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 187360
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-187360

  Query Match             4.4%;  Score 327.7;  DB 1;  Length 440;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  383;  Conservative    7;  Mismatches   40;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
             || |||||||||||| |||||  |||||| | |||| :|||||||||||||||||||| ||
Db        439 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGYGTTAATGACATCAAAGGCATCC 381

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
             |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db        380 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 321

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
             |||| ||  |||||||||||||||||||||||| |||||||||||||| |||||||
Db        320 AGAGTGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 261

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
             ||| |||||||||||||||||||||||||||||||:|||| |||:||||||||:|||||
Db        260 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 201

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
             ||||||||  |||||||||||||| |||||||| |||||||:||||| |  |||||||
Db        200 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 141

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                     |||||||||||||||||||||||||||||| |||||||||  ||||||||||
Db        140 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 81

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||| | :|||| |||  |  |||||||||||||||||||||||||||||| |
Db         80 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 22

Qy       7487 YTTCGAGCCGGCAACGGCAAC 7507
              ||| |||||||||||| ||
Db         21 ATTCAAGCCGGCAACGGCTAC 1


RESULT 538
US-09-925-065A-187360/c
; Sequence 187360, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 187360
;   LENGTH: 440
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-187360

  Query Match             4.4%;  Score 327.7;  DB 1;  Length 440;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  383;  Conservative    7;  Mismatches   40;  Indels   11;  Gaps    3;

Qy       7076 GCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCC 7135
              || |||||||||||| |||||  |||||| | |||| :||||||||||||||||||| ||
Db        439 GCGGACCCCTGGACCAGCCTG-GAGCCCATGCTCTGGYGTTAATGACATCAAAGGCATCC 381

Qy       7136 CTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTT 7195
              |||| |||||||||||| ||||| |||| |||||| ||||||||||||||||||||||||
Db        380 CTCCCGAGGAAATCTCAACTGCATAACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTT 321

Qy       7196 AGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGAC 7255
              |||| ||   |||||||||||||||||||||||| |||||||||||||||| |||||||
Db        320 AGAGTGGATATCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGAC 261

Qy       7256 TGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGA 7315
              ||| |||||||||||||||||||||||||||:|||| ||||:|||||||||:|||||
Db        260 TGACAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGTATCCCTAAGCCTAGCTGGGA 201

Qy       7316 AGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC- 7374
              ||||||||  |||||||||||||| |||||||| |||||||:||||| |  |||||||
Db        200 AGGTGACCGTATCCACCTTTAAACATGGGGCTTGTAACTTAGCTCACATCCAACCAATCA 141

Qy       7375 --------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
                      |||||||||||||||||||||||||||||| ||||||||| |||||||||
Db        140 GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGC 81

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:||| | :|||| |||  |  |||||||||||||||||||||||||| |
Db         80 CAATCATCTATCGTCTGAGGGCA-TGGGGGAGGGACAATGATCGGGATATAAACCCAGGC 22

Qy       7487 YTTCGAGCCGGCAACGGCAAC 7507
              ||| |||||||||||| ||
Db         21 ATTCAAGCCGGCAACGGCTAC 1


RESULT 539
US-09-925-065A-870086
; Sequence 870086, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870086
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-870086

  Query Match             4.4%;  Score 326.7;  DB 1;  Length 448;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  378;  Conservative    2;  Mismatches   37;  Indels    3;  Gaps    3;

Qy       2207 GGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGG 2266
              |||  | || ||| | |||||||||||||| ||:|||||||||||||||| || ||||||
Db         12 GGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGG 69

Qy       2267 GAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAG 2326
              || |||||| |||||||||||||||||||||||| |||||||||||| |||||||||
Db         70 GAACTTTATACTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAG 129

Qy       2327 GGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACAC 2386
              ||||||||||||||| |||||||||||||||||||||||||||:||||||||| ||||||
Db        130 GGAGAGACATTCTAGCGAAAGCAGGGGCCATTATACACCTGAAYATAGGAGAAAGAACAC 189

Qy       2387 CCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 2446
              |||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CCGTTTATTGTCCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 249

Qy       2447 CAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTC 2506
              ||| ||||||||| |||||||||||| | |||||  |||||||||||||||| ||| ||
Db        250 CAACATGGACGAG-CAAAGAATGCCCATTCTGTTTGAGTTAAACTAAAGGATTTCACCTC 308

Qy       2507 CTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGAT 2566
              |||||||||||||||||||||||||||| | |||||||||||||||||| || |||| ||
Db        309 CTTTCCCTACCAAAGGCAGTACCCCCTTACACCCAAGGCCCAACAAGGACTCTAAAATAT 368

Qy       2567 TGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTC 2626
              | |||||||| |||||||||||||| ||||||||||||||  |  ||||||| || | |
Db        369 TTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC 428


RESULT 540
US-09-925-065A-870086
; Sequence 870086, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870086
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-870086

  Query Match             4.4%;  Score 326.7;  DB 1;  Length 448;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  378;  Conservative    2;  Mismatches   37;  Indels    3;  Gaps    3;

Qy       2207 GGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGG 2266
              |||  | || ||| | |||||||||||||| ||:|||||||||||||||| || ||||||
Db         12 GGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGG 69

Qy       2267 GAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAG 2326
              || |||||| |||||||||||||||||||||||| |||||||||||| |||||||||
Db         70 GAACTTTATACTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAG 129

Qy       2327 GGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACAC 2386
              ||||||||||||||| |||||||||||||||||||||||||||:||||||||| ||||||
```

```
Db        130 GGAGAGACATTCTAGCGAAAGCAGGGGCCATTATACACCTGAAYATAGGAGAAAGAACAC 189

Qy       2387 CCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 2446

Db        190 CCGTTTATTGTCCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGA 249

Qy       2447 CAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTC 2506

Db        250 CAACATGGACGAG-CAAAGAATGCCCATTCTGTTTGAGTTAAACTAAAGGATTTCACCTC 308

Qy       2507 CTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGAT 2566

Db        309 CTTTCCCTACCAAAGGCAGTACCCCCTTACACCCAAGGCCCAACAAGGACTCTAAAATAT 368

Qy       2567 TGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTC 2626

Db        369 TTTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC 428


RESULT 541
US-09-864-761-4153/c
; Sequence 4153, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4153
;   LENGTH: 409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF121782.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
US-09-864-761-4153

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169

Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229

Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289

Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349

Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409

Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469

Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519

Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 542
US-10-301-480-393751/c
; Sequence 393751, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 393751
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-393751

  Query Match             4.3%;  Score 324.2;  DB 1;  Length 452;
```

```
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  402;  Conservative    2;  Mismatches   40;  Indels   10;  Gaps    5;

Qy       1698 ATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGA 1757
             ||||||||| | |||:|| |||||||||||||||||||||||||| |||| | || ||||
Db        452 ATTGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGA 393

Qy       1758 CAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTG 1817
             |||||||||||| |||||||||||   ||||||||| ||||||| ||||||||||
Db        392 AAAACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTG 336

Qy       1818 GAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGT 1877
             |||||||||||| |||||||| ||||||| |||||||||||||||| |||| ||||
Db        335 AAGGCTCTGGAAATGGGAAAAGCCGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGT 276

Qy       1878 GCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTC 1937
             ||| |||||||||||||||||||||||||||||||| || |||| |||||| | |||
Db        275 GCGATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTC 216

Qy       1938 CATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTT 1997
             |||||||||||| ||| |||||||||||||||||||||||||||| ||| |||||| |
Db        215 CATGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCT 156

Qy       1998 TTGAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGC 2054
               | |||||||||||||||||||| ||||||||  |||||||||||||||||||||||
Db        155 CTTAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGC 96

Qy       2055 AAGCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGC 2113
               |||| ||| |||||||| |||||||||||||||||  |||||||||||||| ||||||
Db         95 AAGCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGC 36

Qy       2114 CAGGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
              :|||| | |    | ||||||||||||||||||||
Db         35 YAGGAGGTTAACTATCTCCTGGACACTGGTGCGG 2

RESULT 543
US-10-301-480-1007160/c
; Sequence 1007160, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1007160
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1007160

  Query Match             4.3%;  Score 324.2;  DB 1;  Length 452;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  402;  Conservative    2;  Mismatches   40;  Indels   10;  Gaps    5;

Qy       1698 ATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGA 1757
             ||||||||| | |||:|| |||||||||||||||||||||||||| |||| | || ||||
Db        452 ATTGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGA 393

Qy       1758 CAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTG 1817
             |||||||||||| |||||||||||   ||||||||| ||||||| ||||||||||
Db        392 AAAACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTG 336

Qy       1818 GAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGT 1877
             |||||||||||| |||||||| ||||||| |||||||||||||||| |||| ||||

Db        335 AAGGCTCTGGAAATGGGAAAAGCCGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGT 276

Qy       1878 GCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTC 1937
             ||| |||||||||||||||||||||||||||||||| || |||| |||||| | |||
Db        275 GCGATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTC 216

Qy       1938 CATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTT 1997
             |||||||||||| ||| |||||||||||||||||||||||||||| ||| |||||| |
Db        215 CATGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCT 156

Qy       1998 TTGAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGC 2054
               | |||||||||||||||||||| ||||||||  |||||||||||||||||||||||
Db        155 CTTAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGC 96

Qy       2055 AAGCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGC 2113
               |||| ||| |||||||| |||||||||||||||||  |||||||||||||| ||||||
Db         95 AAGCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGC 36

Qy       2114 CAGGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
              :|||| | |    | ||||||||||||||||||||
Db         35 YAGGAGGTTAACTATCTCCTGGACACTGGTGCGG 2

RESULT 544
US-09-925-065A-399985/c
; Sequence 399985, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 399985
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-399985

  Query Match             4.3%;  Score 323.8;  DB 1;  Length 450;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   39;  Indels   10;  Gaps    5;

Qy       1700 TGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACA 1759
             ||||||| | |||:|| |||||||||||||||||||||||||| |||| | || |||| |
Db    .   450 TGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAA 391

Qy       1760 AACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGA 1819
             |||||||||| |||||||||||   ||||||||| ||||||| ||||||||||| |
Db        390 AACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAA 334

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
             |||||||||| |||||||||:|||||||| |||||||||||||||| |||| ||||||
Db        333 GGCTCTGGAAATGGGAAAAGCYGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGC 274

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
             | |||||||||||||||||||||||||||||||||| || |||| |||||| | |||||
Db        273 GATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCA 214

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
```

```
Db        213 TGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCT 154

Qy       2000 GAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAA 2056
Db        153 TAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAA 94

Qy       2057 GCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCA 2115
Db         93 GCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCCA 34

Qy       2116 GGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
Db         33 GGAGGTTAACTATCTCCTGGACACTGGTGCGG 2

RESULT 545
US-09-925-065A-399985/c
; Sequence 399985, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 399985
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-399985

  Query Match             4.3%;  Score 323.8;  DB 1;  Length 450;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   39;  Indels   10;  Gaps    5;

Qy       1700 TGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACA 1759
Db        450 TGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAA 391

Qy       1760 AACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGA 1819
Db        390 AACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAA 334

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
Db        333 GGCTCTGGAAATGGGAAAAGCYGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGC 274

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
Db        273 GATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCA 214

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
Db        213 TGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCT 154

Qy       2000 GAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAA 2056
Db        153 TAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAA 94

Qy       2057 GCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCA 2115
Db         93 GCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCCA 34

Qy       2116 GGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
Db         33 GGAGGTTAACTATCTCCTGGACACTGGTGCGG 2

RESULT 546
US-09-925-065A-318955/c
; Sequence 318955, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 318955
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-318955

  Query Match             4.3%;  Score 322.2;  DB 1;  Length 450;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  400;  Conservative    2;  Mismatches   40;  Indels   10;  Gaps    5;

Qy       1700 TGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACA 1759
Db        450 TGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAA 391

Qy       1760 AACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGA 1819
Db        390 AACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAA 334

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
Db        333 GGCTCTGGAAATGGGAAAAGCCGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGC 274

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
Db        273 GATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCA 214

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
Db        213 TGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCT 154

Qy       2000 GAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAA 2056
Db        153 TAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAA 94

Qy       2057 GCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCA 2115
Db         93 GCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCYA 34

Qy       2116 GGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
Db         33 GGAGGTTAACTATCTCCTGGACACTGGTGCGG 2
```

```
RESULT 547
US-09-925-065A-318955/c
; Sequence 318955, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 318955
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-318955

  Query Match             4.3%;  Score 322.2;  DB 1;  Length 450;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  400;  Conservative    2;  Mismatches   40;  Indels   10;  Gaps    5;

Qy       1700 TGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACA 1759
             ||||||| | |||:|| |||||||||||||||||||||||||| |||| | || |||| |
Db        450 TGAACTTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAA 391

Qy       1760 AACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGA 1819
             |||||||||| |||||||||||   |||||||| ||||||||  |||||||||| |
Db        390 AACGGGATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAA 334

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
             |||||||||| ||||||||| |||||||| |||||||||||||||||| |||| ||||||
Db        333 GGCTCTGGAAATGGGAAAAGCCGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGC 274

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
             | |||||||||||||||||||||||||||||||||| || ||||  ||||| | |||||
Db        273 GATCTACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCA 214

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
             |||||||||| ||| |||||||||||||||||||||||||||||| ||| |||||| | |
Db        213 TGCCCCTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCT 154

Qy       2000 GAGTCAGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAA 2056
              |||||||||||||||||||| ||||||||  |||||||||||||||||||||||||||
Db        153 TAGTCAGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAA 94

Qy       2057 GCGCCATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCA 2115
              || ||| |||||||| |||||||||||||||||  |||||||||||||| |||||||:|
Db         93 GCACCAGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCYA 34

Qy       2116 GGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
              ||| | |   | ||||||||||||||||||||
Db         33 GGAGGTTAACTATCTCCTGGACACTGGTGCGG 2

RESULT 548
US-09-925-065A-863188
; Sequence 863188, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863188
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-863188

  Query Match             4.3%;  Score 322;  DB 1;  Length 452;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  385;  Conservative    2;  Mismatches   38;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
             | | |  | |||  | || ||| | |||||||||||||| ||:|||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
             || | |||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db         62 TG-AGTGGGGAGCTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
             |||||||||||||| ||||||||||||||||||||||||||||||:|||||||||||||
Db        121 CCTTGTTAGGGAGACACATTCTAGCAAAAGCAGGGGCCATTATACAMCTGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
             |||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCTGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
             ||||||||||||||||||| || |||||||||||| ||||||||||||||||||||||||
Db        240 ACAGAAGGACAATATGGACCAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
             |||| | | |||||||||||| |||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCGCCTTCTTTCCCTACCGAAGGCAGTACCCCCTTAGACCGGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
             |||||||||||||||||| |||||||| ||||| |||||||||||||||   | |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCTAAGGCCTAGTAAAACCATGCAATAGCCCCTG 418

Qy       2618 CAGTAATTC 2626
              || || | |
Db         419 CAATACTCC 427

RESULT 549
US-09-925-065A-863188
; Sequence 863188, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863188
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-863188

  Query Match             4.3%;  Score 322;  DB 1;  Length 452;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  385;  Conservative    2;  Mismatches   38;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||||||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || | |||||||||||| |||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-AGTGGGGAGCTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
                |||||||||||||| |||||||||||||||||||||||||||||:||||||||||||||
Db        121 CCTTGTTAGGGAGACACATTCTAGCAAAAGCAGGGGCCATTATACAMCTGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCTGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                |||||||||||||||||| || |||||||||||| ||||||||||||||||||||||||||
Db        240 ACAGAAGGACAATATGGACCAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                |||| | | |||||||||||| |||||||||||||| |||||  ||||||||||||| |
Db        299 TTCCGCCTTCTTTCCCTACCGAAGGCAGTACCCCCTTAGACCGGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                |||||||||||||||||| |||||||| ||||| |||||||||||||   | |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCTAAGGCCTAGTAAAACCATGCAATAGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 550
US-09-925-065A-890779/c
; Sequence 890779, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890779
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-890779

  Query Match             4.3%;  Score 321.5;  DB 1;  Length 449;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   42;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
                ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
                |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
                |||||||||||||||||||||||||| |||||||||||||||||||||| | ||||:||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGYTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
                ||| |||||||||||| |||||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
                |||||||||||||| |||||||| ||||||   |  ||||||| || | |       |
Db          93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                |   |  | |||||||||||   |||||||||| |
Db          40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 551
US-09-925-065A-890779/c
; Sequence 890779, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890779
;   LENGTH: 449
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-925-065A-890779

  Query Match             4.3%;  Score 321.5;  DB 1;  Length 449;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   42;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
              ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
              |||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
              |||||||||||||||||||||||| |||||||||||||||||||| | ||||:||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGYTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
              ||| |||||||||||| |||||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
              |||||||||||||| |||||||| ||||||   |  ||||||| || | |         |
Db          93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
              |    |   | ||||||||||||   |||||||||||| |
Db          40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 552
US-09-925-065A-914003/c
; Sequence 914003, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 914003
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-914003

  Query Match             4.3%;  Score 321.5;  DB 1;  Length 449;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   42;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
              ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
              |||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
              |||||||||||||||||||||||| |||||||||||||||||||| | ||||:||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
              ||| |||||||||||| |||||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
              |||||||||||||| |||||||| ||||||:|||||  |  ||||||| || | |         |
Db          93 CTAAAAGCCCAAGGCCTAGTAAAASCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
              |    |   | ||||||||||||   |||||||||||| |
Db          40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 553
US-09-925-065A-914003/c
; Sequence 914003, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 914003
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-914003

  Query Match             4.3%;  Score 321.5;  DB 1;  Length 449;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  401;  Conservative    2;  Mismatches   42;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
              ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
              |||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332
```

```
Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
             |||||||||||||||||||||||||||| |||||||||||||||||||||||| | |||| ||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
             ||| |||||||||||| |||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
              |||||||||||||| ||||||||:|||||| | |||||| || | |        |
Db         93 CTAAAAGCCCAAGGCCTAGTAAAASCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
             | | | |||||||||||   ||||||||||| |
Db         40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5
```

RESULT 554
US-09-925-065A-914004/c
; Sequence 914004, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 914004
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-914004

Query Match 4.3%; Score 321.5; DB 1; Length 449;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 401; Conservative 2; Mismatches 42; Indels 11; Gaps 5;

```
Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
             ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
             |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
             |||||||||||||||||||||||||||| |||||||||||||||||||||||| | |||| ||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
             ||| |||||||||||| |||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCRAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
              |||||||||||||| |||||||| |||||| | ||||||| || | |        |
Db         93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
             | | | |||||||||||   ||||||||||| |
Db         40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5
```

RESULT 555
US-09-925-065A-914004/c
; Sequence 914004, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 914004
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-914004

Query Match 4.3%; Score 321.5; DB 1; Length 449;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 401; Conservative 2; Mismatches 42; Indels 11; Gaps 5;

```
Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
             ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
             |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
             |||||||||||||||||||||||||||| |||||||||||||||||||||||| | |||| ||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
             ||| |||||||||||| |||||||||||||||||||||||||| | ||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
```

```
                ||||  ||||||||||||  ||||||:||  ||||||||  ||  ||||||||||||  ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCRAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
                |||||||||||||||  ||||||||  ||||||   |  |||||||  ||  | |         |
Db         93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                |   |  | ||||||||||||   ||||||||||| |
Db          40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 556
US-09-925-065A-893796
; Sequence 893796, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893796
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893796

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 441;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || |||||||| || || |||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
                ||||||||||||||||||||||||||||||||||:||||||||||| ||||||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGYCATTATACACCCGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                ||||||  ||  ||||  |||  |||  |||||||||||  ||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                ||||||  ||||||||||||||||||||||||||||||  |||||  |||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                ||||||||||||||||||  ||||||||||||  ||||||||||||||||    |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 557
US-09-925-065A-893796
; Sequence 893796, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893796
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893796

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 441;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || |||||||| || || |||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
                ||||||||||||||||||||||||||||||||||:||||||||||| ||||||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGYCATTATACACCCGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                ||||||  ||  ||||  |||  |||  |||||||||||  ||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                ||||||  ||||||||||||||||||||||||||||||  |||||  |||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                ||||||||||||||||||  ||||||||||||  ||||||||||||||||    |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 558
US-09-925-065A-893797
```

```
; Sequence 893797, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893797
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893797

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 441;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || |||||||| || || ||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCTGAACATAGGAG 2377
                |||||||||||||||||||||||||||||||||||| ||||||||||:||||||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCYGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                |||||| || |||| ||| ||| |||||||||||| ||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                |||||| |||||||||||||||||||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                |||||||||||||||||| |||||||||||||| ||||||||||||||||     |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 559
US-09-925-065A-893797
; Sequence 893797, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893797
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893797

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 441;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || |||||||| || || ||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCYGAACATAGGAG 2377
                |||||||||||||||||||||||||||||||||||| ||||||||||:||||||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCYGAACATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                |||||| || |||| ||| ||| |||||||||||| ||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                |||||| |||||||||||||||||||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                |||||||||||||||||| |||||||||||||| ||||||||||||||||     |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 560
US-09-925-065A-863189
; Sequence 863189, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863189
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-863189

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 452;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
             | | |  | |||  | || ||| | |||||||||||||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
             || | |||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db         62 TG-AGTGGGGAGCTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
             ||||||||||||||| |||||||||||||||||||||||||||||||| |||||:|||||||
Db        121 CCTTGTTAGGGAGACACATTCTAGCAAAAGCAGGGGCCATTATACAACTGAAYATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
             |||||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCTGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
             |||||||||||||||||| || |||||||||||| ||||||||||||||||||||||||||
Db        240 ACAGAAGGACAATATGGACCAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
             |||| | | |||||||||||| |||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCGCCTTCTTTCCCTACCGAAGGCAGTACCCCCTTAGACCGGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
             |||||||||||||||||| |||||||| ||||| |||||||||||||||  |  |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCTAAGGCCTAGTAAAACCATGCAATAGCCCCTG 418

Qy       2618 CAGTAATTC 2626
             || || | |
Db        419 CAATACTCC 427


RESULT 561
US-09-925-065A-863189
; Sequence 863189, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863189
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-863189

  Query Match             4.3%;  Score 320.4;  DB 1;  Length 452;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  384;  Conservative    2;  Mismatches   39;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
             | | |  | |||  | || ||| | |||||||||||||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
             || | |||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db         62 TG-AGTGGGGAGCTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
             ||||||||||||||| |||||||||||||||||||||||||||||||| |||||:|||||||
Db        121 CCTTGTTAGGGAGACACATTCTAGCAAAAGCAGGGGCCATTATACAACTGAAYATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
             |||||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCTGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
             |||||||||||||||||| || |||||||||||| ||||||||||||||||||||||||||
Db        240 ACAGAAGGACAATATGGACCAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
             |||| | | |||||||||||| |||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCGCCTTCTTTCCCTACCGAAGGCAGTACCCCCTTAGACCGGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
             |||||||||||||||||| |||||||| ||||| |||||||||||||||  |  |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCTAAGGCCTAGTAAAACCATGCAATAGCCCCTG 418

Qy       2618 CAGTAATTC 2626
             || || | |
Db        419 CAATACTCC 427


RESULT 562
US-09-925-065A-890778/c
; Sequence 890778, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890778
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-890778

  Query Match             4.3%;  Score 319.9;  DB 1;  Length 449;
```

```
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  400;  Conservative    2;  Mismatches   43;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
             ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
             |||||||||||||||||||||||||||| |||||||||||||||||||||||| | |||| ||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
             | ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTRGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
             ||| |||||||||||| |||||||||||||||||||||||||||| | ||||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
             ||||||||||||||| |||||||| ||||||  |  ||||||| || | |         |
Db         93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
             |   |  | ||||||||||||   |||||||||||  |
Db         40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 563
US-09-925-065A-890778/c
; Sequence 890778, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 890778
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-890778

  Query Match             4.3%;  Score 319.9;  DB 1;  Length 449;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  400;  Conservative    2;  Mismatches   43;  Indels   11;  Gaps    5;

Qy       2217 GACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATT 2276
             ||| |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |
Db        449 GACAGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACT 392

Qy       2277 CTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACAT 2336
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db        391 CTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACAT 332

Qy       2337 TCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTG 2396
             |||||||||||||||||||||||||||| |||||||||||||||||||||||| | |||| ||
Db        331 TCTAGCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTG 272

Qy       2397 TNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGAC 2456
             | ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        271 T-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTRGGCAACAGAAGGACAATATGGAC 213

Qy       2457 GAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTAC 2516
             ||| |||||||||||| |||||||||||||||||||||||||||| | ||||||||||||
Db        212 GAG-CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAA 154

Qy       2517 CAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||| |||||||||||| |||||| || |||||||| || |||||||||| ||||||||
Db        153 CAAAAGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGAC 94

Qy       2577 TTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGAT 2636
             ||||||||||||||| |||||||| ||||||  |  ||||||| || | |         |
Db         93 CTAAAAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAA 41

Qy       2637 TGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
             |   |  | ||||||||||||   |||||||||||  |
Db         40 TTTAGGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 564
US-09-925-065A-879842/c
; Sequence 879842, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 879842
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-879842

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
             |||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |||||
Db        445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
             |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db        387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
             |||||||||||||||||||| |||||||||||||||||||||||||||| | ||||:||| ||
Db        327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGYTGT-CC 269
```

```
Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
             ||||||||||||  |||||||||||||||||||||||||||||| | ||||||||||| ||||
Db        209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
              ||||||||||| |||||| || |||||||| || |||||||||| |||||||| |||
Db        149 AGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
             ||||||||||| |||||||| ||||||  |  ||||||| || | |        | |
Db         89 AAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
              |  | ||||||||||   ||||||||||| |
Db         36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 565
US-09-925-065A-879842/c
; Sequence 879842, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 879842
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-879842

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
             ||||||||||||||||  ||:|||||||||||||||| || |||||||| ||||| |||||
Db        445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
             |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db        387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
             |||||||||||||||||||| ||||||||||||||||||||| | ||||:||| ||
Db        327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGYTGT-CC 269

Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
             ||||||||||||  |||||||||||||||||||||||||||||| | ||||||||||  ||||
Db        209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
              ||||||||||| |||||| || |||||||| || |||||||||| |||||||| |||
Db        149 AGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
             ||||||||||| |||||||| ||||||  |  ||||||| || | |        | |
Db         89 AAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
              |  | ||||||||||   ||||||||||| |
Db         36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 566
US-09-925-065A-907856/c
; Sequence 907856, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 907856
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-907856

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
             ||||||||||||||||  ||:|||||||||||||||| || |||||||| ||||| |||||
Db        445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
             |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db        387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
             |||||||||||||||||||| ||||||||||||||||||||| | ||||:||| ||
Db        327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTGT-CC 269

Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
             ||||||||||||  |||||||||||||||||||||||||||||| | ||||||||||  ||||
Db        209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
              ||||||||||| |||||| || |||||||| || |||||||||| |||||||| |||
Db        149 AGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
```

```
                ||||||||||| |||||||:|||||| | ||||||| || | |       | |
Db         89 AAGCCCAAGGCCTAGTAAAASCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                | | ||||||||||| ||||||||||| |
Db         36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 567
US-09-925-065A-907856/c
; Sequence 907856, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 907856
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-907856

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
                ||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |||||
Db         445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
                |||||||||||||||||||| ||||||||||||| ||||||||||||||||||||||||
Db         387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
                |||||||||||||||||||| |||||||||||||||||||||||| | |||| ||| ||
Db         327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTGT-CC 269

Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
                |||||||||||| |||||||||||||||||||||||||||| | |||||||||||| ||||
Db         209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
                |||||||||||| |||||| || |||||||| || |||||||||| |||||||| |||
Db         149 AGCAGTACCCCCTTAGACCCGAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
                |||||||||||| ||||||||:|||||| | ||||||| || | |       | |
Db          89 AAGCCCAAGGCCTAGTAAAASCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                |  | ||||||||||| ||||||||||| |
Db          36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 568
US-09-925-065A-907857/c
; Sequence 907857, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 907857
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-907857

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
                ||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |||||
Db         445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
                |||||||||||||||||||| ||||||||||||| ||||||||||||||||||||||||
Db         387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
                |||||||||||||||||||| |||||||||||||||||||||||| | |||| ||| ||
Db         327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTGT-CC 269

Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
                |||||||||||| |||||||||||||||||||||||||||| | |||||||||||| ||||
Db         209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
                |||||||||||| |||||| || |||||||| || |||||||||| |||||||| |||
Db         149 AGCAGTACCCCCTTAGACCCRAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
                |||||||||||| ||||||||:|||||| | ||||||| || | |       | |
Db          89 AAGCCCAAGGCCTAGTAAAASCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                |  | ||||||||||| ||||||||||| |
Db          36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 569
US-09-925-065A-907857/c
; Sequence 907857, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 907857
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-907857

  Query Match             4.3%;  Score 319.1;  DB 1;  Length 446;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  398;  Conservative    2;  Mismatches   41;  Indels   11;  Gaps    5;

Qy       2221 GGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTT 2280
                ||||||||||||||||| ||:|||||||||||||||| || |||||||| ||||| |||||
Db        445 GGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTGTG-ACTGGGGAACTTTACTCTTT 388

Qy       2281 TCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTA 2340
                |||||||||||||||||||| |||||||||||||| ||||||||||||||||||||||
Db        387 TCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTCTA 328

Qy       2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCC 2400
                |||||||||||||||||||| |||||||||||||||||||||| | |||| ||| ||
Db        327 GCAAAAGCAGGGGCCATTATACCCCTGAACATAGGAGAAGGAACACACATTTGCTGT-CC 269

Qy       2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGC 2460
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG- 210

Qy       2461 CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAA 2520
                |||||||||||| |||||||||||||||||||||||| | |||||||||| ||||
Db        209 CAAAGAATGCCCATCCTGTTCAAGTTAAACTAAAGGATTCCGCCTCCTTTCCCTAACAAA 150

Qy       2521 GGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAA 2580
                 |||||||||||| ||||||:|| |||||||| || |||||||||| |||||||| |||
Db        149 AGCAGTACCCCCTTAGACCCRAGCCCCAACAATGACTCCAAAAGATTTTTAAGGACCTAA 90

Qy       2581 AAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAG 2640
                |||||||||| |||||||| ||||||  |  ||||||| || | |        | |
Db         89 AAGCCCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGCAATACTCC-------AAATTTA 37

Qy       2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGT 2672
                |  | ||||||||||    |||||||||| |
Db         36 GGAGTACAGAAACCCATCAGACAGTGGAGGTT 5


RESULT 570
US-09-925-065A-893798
; Sequence 893798, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893798
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893798

  Query Match             4.3%;  Score 318.8;  DB 1;  Length 441;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  383;  Conservative    2;  Mismatches   40;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
                | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
                || |||||||| || || |||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
                |||||||||||||||||||||||||||||||||||| |||||||||| |||:|||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCCGAAYATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
                |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
                |||||| || |||| ||| ||| |||||||||| ||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
                |||||| |||||||||||||||||||||||||| |||||  |||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
                |||||||||||||||||| |||||||||||||| ||||||||||||||||     |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
                || || | |
Db        419 CAATACTCC 427


RESULT 571
US-09-925-065A-893798
; Sequence 893798, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 893798
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-893798

  Query Match             4.3%;  Score 318.8;  DB 1;  Length 441;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  383;  Conservative    2;  Mismatches   40;  Indels    4;  Gaps    4;

Qy       2198 ATTCTGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTA 2257
              | | |  | |||  | || ||| | |||||||||| ||| ||:||||||||||||||||
Db          3 ACTATCCGAGGGGTCCTAGGACAGCCAGTCACTAGCTAC-TTCTCCCAGCCACTAAGTTG 61

Qy       2258 TGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTA 2317
              || |||||||| || || |||||||||||||||||||||||||||| ||||||||||||
Db         62 TG-ACTGGGGAACTCTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTC 120

Qy       2318 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAG 2377
              |||||||||||||||||||||||||||||||||||||||| |||||||||| |||:|||||||
Db        121 CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGTCATTATACACCCGAAYATAGGAG 180

Qy       2378 AAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 2437
              |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAACACCCATTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCA 239

Qy       2438 ACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGA 2497
              |||||| || |||| ||| ||| |||||||||| |||||||||||||||||||||||||
Db        240 ACAGAAAGAAAATAAGGAAGAG-CAAAGAATGCCTGTCCTGTTCAAGTTAAACTAAAGGA 298

Qy       2498 TTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATT 2557
              |||||| |||||||||||||||||||||||||||||| |||||  ||||||||||||| |
Db        299 TTCCACCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACT 358

Qy       2558 CCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTG 2617
              |||||||||||||||||| |||||||||||||| |||||||||||||||||     |||||
Db        359 CCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCATTCGCCCCTG 418

Qy       2618 CAGTAATTC 2626
              || || | |
Db        419 CAATACTCC 427


RESULT 572
US-10-301-480-467328/c
; Sequence 467328, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 467328
;   LENGTH: 445
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-467328

  Query Match             4.3%;  Score 318.8;  DB 1;  Length 445;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  396;  Conservative    2;  Mismatches   39;  Indels   10;  Gaps    5;

Qy       1705 TTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGG 1764
              || | |||:|| |||||||||||||||||||||||||||| |||| | || |||| ||||||
Db         445 TTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAAAACGG 386

Qy       1765 GATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTC 1824
              |||||| ||||||||||||   ||||||||| |||||||| |||||||||| ||||||
Db         385 GATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAAGGCTC 329

Qy       1825 TGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCT 1884
              |||||| |||||||||:|||||||| |||||||||||||||||| |||| ||||||| |||
Db         328 TGGAAATGGGAAAAGCYGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGCGATCT 269

Qy       1885 ACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCC 1944
              |||||||||||||||||||||||||||||||||||| || |||| ||||| | ||||||||||
Db         268 ACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCATGCCC 209

Qy       1945 CTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTC 2004
              ||||| ||| |||||||||||||||||||||||||||||| ||| |||||| | | ||||
Db         208 CTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCTTAGTC 149

Qy       2005 AGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAAGCGCC 2061
              |||||||||||||||| |||||||||  |||||||||||||||||||||||||||||| ||
Db         148 AGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAAGCACC 89

Qy       2062 ATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG 2120
              | |||||||| |||||||||||||||||||||||| |||||||||||||| |||||||| |
Db          88 AGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCCAGGAGG 29

Qy       2121 GT---TGTCTCCTGGACACTGGTGCGG 2144
              ||   | |||||||||||||||||||||
Db          28 TTAACTATCTCCTGGACACTGGTGCGG 2


RESULT 573
US-10-301-480-1080737/c
; Sequence 1080737, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1080737
;   LENGTH: 445
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1080737

  Query Match             4.3%;  Score 318.8;  DB 1;  Length 445;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  396;  Conservative    2;  Mismatches   39;  Indels   10;  Gaps    5;

Qy       1705 TTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGG 1764
              || | |||:|| |||||||||||||||||||||||||||| |||| | || |||| ||||||
Db         445 TTAGGAACCTCAGTTTTTTATAATAGAGATCAGGAGGAACAGGTGCAAAAGGAAAAACGG 386

Qy       1765 GATTAAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTC 1824
              |||||| ||||||||||||   ||||||||| |||||||| |||||||||| ||||||
Db         385 GATTAACAAAAAGGCCACC---TTAGTCATGGCCCTCAGGTGAGTGGACTTTGAAGGCTC 329

Qy       1825 TGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCT 1884
              |||||| |||||||||:|||||||| |||||||||||||||||| |||| ||||||| |||
Db         328 TGGAAATGGGAAAAGCYGGGCAAATAGAATGCCTAATAGGGCTCGCTTTCAGTGCGATCT 269

Qy       1885 ACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCC 1944
```

```
                ||||||||||||||||||||||||||||||| || ||||  |||||  | ||||||||||
Db        268 ACAAGGACACTTTAAAAAAGATTGTCCAAGTATAAATAAGTTGCCCCCTTGTCCATGCCC 209

Qy       1945 CTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTC 2004
              |||||  ||| |||||||||||||||||||||||||||||||| ||| ||||||  | | ||||
Db        208 CTTATGTCAGGGGAATCACTGGAAGGCCCACTGCCCCAGGAGACGAAGGTCCTCTTAGTC 149

Qy       2005 AGAAGCCACTAACCAGA-TGATCCAGC--AGCAGGACTGAGGGTGCCTGGGGCAAGCGCC 2061
              |||||||||||||||||  |||||||||  |||||||||||||||||||||||||||| ||
Db        148 AGAAGCCACTAACCAGATTGATCCAGCAGAGCAGGACTGAGGGTGCCTGGGGCAAGCACC 89

Qy       2062 ATCCCATGCC-ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAG 2120
              | ||||||||  ||||||||||||||||||  |||||||||||||| |||||||||| |
Db         88 AGCCCATGCCAATCACCCTCACAGAGCCCCAGGTATGCTTGACCATGGAGGGCCAGGAGG 29

Qy       2121 GT---TGTCTCCTGGACACTGGTGCGG 2144
                |   | ||||||||||||||||||||
Db         28 TTAACTATCTCCTGGACACTGGTGCGG 2
```

RESULT 574
US-09-864-761-4227/c
; Sequence 4227, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4227
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034410.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
US-09-864-761-4227

Query Match 4.2%; Score 315.5; DB 1; Length 414;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 347; Conservative 9; Mismatches 44; Indels 1; Gaps 1;

```
Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
              ||||||  |||||  ||| |||||||||||||||||||| ||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
              |||||||  |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              |||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
              ||||||| ||  |||||||| |||||| ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
              |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              ||||||  |||||  ||||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15
```

RESULT 575
US-09-925-065A-147128/c
; Sequence 147128, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 147128
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-147128

  Query Match             4.1%;  Score 309.6;  DB 1;  Length 440;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  366;  Conservative    6;  Mismatches   68;  Indels    2;  Gaps    2;

Qy        120 ACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTG 179
              ||||||| ||||| ||||| ||| ||||| |||||||||||| |||| || |||||||:||
Db        440 ACTCTTCAGGTCCGTGTTTGTTATGGCTCAAGCTGAGCTTTCGCTCGCCATCCACCAMTG 381

Qy        180 CTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGT 239
              ||||||||| ||  | || ||| |||| ||||| ||| || ||||||||| |||||||||
Db        380 CTGTTTGCCGCCGTCCCAGACCCGCCGTTGACTTCCACCCTTCTGGATCCAGCAGGGTGT 321

Qy        240 CCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTG 299
              ||  ||  ||||||||||||||:| ||||||||||||||  ||||||||||||||||
Db        320 CCATTGCACTCCTGATCCAGCGAGGTGCCCATTGCCGCTCCTGATTGGGCTAAAGGCTCA 261

Qy        300 CCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCAC 359
              ||||||| ||||||| | |||||||| || ||||  ||| ||||||||||||||| ||
Db        260 CCATTGTTCCTGCACAGGTAAGTGCCCGGCTTTGACCTAGTTGAGCTGAACACTAGTCGG 201

Qy        360 TGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCAC 419
              ||| |||||||||||||||||||||||||| ||||||||||:| |||||||||| |||||
Db        200 TGGCTTCCATGGTTCTCTTCTGTGACCCACAGCTTCTAATAGAGCTATAACACTCACCAC 141

Qy        420 ATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATA 479
              |||||||||| ||| |||||||||||||| ||||||:||| ||||   |||||||||| |
Db        140 ATGGCCCAAGCTTCTATTCCTTGGAATCTGTGAGGCCAA-GAACCTTAGGTCAGAGAACA 82

Qy        480 CGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCAT 539
                |:| ||||| ||||||||| ||| || |||  ||:|||||| ||  |  | |||||||
Db         81 AAAGGTTTGCCGCCATCTTGGGAGCCGCTTGC-CCCATCTTGGGAGCCGCCCGCCACCAT 23

Qy        540 CTTGGGAGCTCTGTGAGCAAGG 561
              ||||||||||||  ||  ||||
Db         22 CTTGGGAGCTCTAAGAATAAGG 1


RESULT 576
US-09-925-065A-147128/c
; Sequence 147128, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 147128
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-147128

  Query Match             4.1%;  Score 309.6;  DB 1;  Length 440;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  366;  Conservative    6;  Mismatches   68;  Indels    2;  Gaps    2;

Qy        120 ACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTG 179
              ||||||| ||||| ||||| ||| ||||| |||||||||||| |||| || |||||||:||
Db        440 ACTCTTCAGGTCCGTGTTTGTTATGGCTCAAGCTGAGCTTTCGCTCGCCATCCACCAMTG 381

Qy        180 CTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGT 239
              ||||||||| ||  | || ||| |||| ||||| ||| || ||||||||| |||||||||
Db        380 CTGTTTGCCGCCGTCCCAGACCCGCCGTTGACTTCCACCCTTCTGGATCCAGCAGGGTGT 321

Qy        240 CCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTG 299
              ||  ||  ||||||||||||||:| ||||||||||||||  ||||||||||||||||
Db        320 CCATTGCACTCCTGATCCAGCGAGGTGCCCATTGCCGCTCCTGATTGGGCTAAAGGCTCA 261

Qy        300 CCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCAC 359
              ||||||| ||||||| | |||||||| || ||||  ||| ||||||||||||||| ||
Db        260 CCATTGTTCCTGCACAGGTAAGTGCCCGGCTTTGACCTAGTTGAGCTGAACACTAGTCGG 201

Qy        360 TGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCAC 419
              ||| |||||||||||||||||||||||||| ||||||||||:| |||||||||| |||||
Db        200 TGGCTTCCATGGTTCTCTTCTGTGACCCACAGCTTCTAATAGAGCTATAACACTCACCAC 141

Qy        420 ATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATA 479
              |||||||||| ||| |||||||||||||| ||||||:||| ||||   |||||||||| |
Db        140 ATGGCCCAAGCTTCTATTCCTTGGAATCTGTGAGGCCAA-GAACCTTAGGTCAGAGAACA 82

Qy        480 CGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCAT 539
                |:| ||||| ||||||||| ||| || |||  ||:|||||| ||  |  | |||||||
Db         81 AAAGGTTTGCCGCCATCTTGGGAGCCGCTTGC-CCCATCTTGGGAGCCGCCCGCCACCAT 23

Qy        540 CTTGGGAGCTCTGTGAGCAAGG 561
              ||||||||||||  ||  ||||
Db         22 CTTGGGAGCTCTAAGAATAAGG 1


RESULT 577
US-10-040-916-47
; Sequence 47, Application US/10040916
; Publication No. US20020146769A1
;   GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John
;                  LaVallie, Edward
;                  Racie, Lisa
;                  Merberg, David
;                  Treacy, Maurice
;                  Evans, Cheryl
;                  Spaulding, Vikki
;       TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
;                           ENCODING THEM
;       NUMBER OF SEQUENCES: 71
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/040,916
;            FILING DATE: 07-Jan-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/887,029
;            FILING DATE: 07-FEB-1997
;            APPLICATION NUMBER: 08/686,878
;            FILING DATE: <Unknown>
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Brown, Scott A.
;            REGISTRATION NUMBER: 32,724
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 498-8224
;            TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 47:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 374 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-10-040-916-47

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
              ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
              |||||||||||||||||||||||||| ||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
              |  ||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
              |  |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              ||||  ||  |||||||||||||||||||||| ||||| |||||| |||| ||||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 578
US-09-864-761-24312/c
; Sequence 24312, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24312
;   LENGTH: 336
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC022555.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;   OTHER INFORMATION: SWISSPROT HIT: P26808, EVALUE 6.00e-18
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130
US-09-864-761-24312

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:||||||||||  |||||||||||||||||:|||||::|||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||: ||||||||||:|||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
              |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 579
US-09-864-761-30194/c
```

```
; Sequence 30194, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30194
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004909.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;   OTHER INFORMATION: EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
;   OTHER INFORMATION: SWISSPROT HIT: P34109, EVALUE 2.50e-01
;   OTHER INFORMATION: NT HIT: gi11419444, EVALUE 3.00e-96
US-09-864-761-30194

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||

Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||

Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||| |||||| |||||||||   |||| || || |

Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||

Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||:|||

Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||| | |||||||| ||| ||| ||:| |

Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 580
US-09-925-065A-817633/c
; Sequence 817633, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817633
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-817633

  Query Match             3.7%;  Score 279.4;  DB 1;  Length 332;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches  286;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       7270 AGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCC 7329
              |||||||||||||||||||:|||||||||:||||||||| |||||||||||| |||||
Db        332 AGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCC 273

Qy       7330 ACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAA 7389
              |||||||||||| | ||||||||||||:|||||||  ||||||||||||||||||||||
Db        272 ACCTTTAAACACAGAGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAA 213

Qy       7390 ATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCA 7449
              || |||||||||||||| |||||| ||||||||||||||||||:||||||:||||||||
Db        212 ATACTAATTAGGCAAAAACAGGAAGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCA 153

Qy       7450 CAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCC 7509
```

```
              ||||||||||||||| |||| ||||||||||||| |  |||||||| ||||| ||||||||
Db        152 CAGCAGGAGGGACAAGGATCAGGATATAAACCCAGGCATTCGAGCTGGCAAGGGCAACCC 93

Qy       7510 CCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAA 7569
              :|||||||||||||||| |||||||||||||||||||||  |||||||||||||||||||
Db         92 SCTTTGGGTCCCCTCCCCTTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAA 33

Qy       7570 TCTTGCARCTGCR 7582
              |||||||:||||:
Db         32 TCTTGCAACTGCA 20


RESULT 581
US-09-925-065A-817633/c
; Sequence 817633, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817633
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-817633

  Query Match             3.7%;  Score 279.4;  DB 1;  Length 332;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches  286;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       7270 AGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCC 7329
              |||||||||||||||||||:|||||||||:|||||||| ||||||||||||| |||||
Db        332 AGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCC 273

Qy       7330 ACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAA 7389
              |||||||||||| | |||||||||||||:|||||||  ||||||||||||||||||||||
Db        272 ACCTTTAAACACAGAGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAA 213

Qy       7390 ATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCA 7449
              || |||||||||||| |||||| ||||||||||||||||||:||||||:||||||||
Db        212 ATACTAATTAGGCAAAAACAGGAAGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCA 153

Qy       7450 CAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCC 7509
              |||||||||||||||| |||| |||||||||||| |  ||||||| ||||| ||||||||
Db        152 CAGCAGGAGGGACAAGGATCAGGATATAAACCCAGGCATTCGAGCTGGCAAGGGCAACCC 93

Qy       7510 CCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAA 7569
              :|||||||||||||||| |||||||||||||||||||||  |||||||||||||||||||
Db         92 SCTTTGGGTCCCCTCCCCTTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAA 33

Qy       7570 TCTTGCARCTGCR 7582
              |||||||:||||:
Db         32 TCTTGCAACTGCA 20


RESULT 582
US-09-864-761-10242/c
; Sequence 10242, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10242
;   LENGTH: 379
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007244.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
US-09-864-761-10242

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
              |||||||||||  ||   ||| |      ||  |||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324
```

```
Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             ||||||||||||| |||| |||||||||||||||||||||| ||||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  ||||||||||||||||||||| ||| |||||||||||  ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |   | |||||||||||||||||||||||| ||||   |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 583
US-09-864-761-18530
; Sequence 18530, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18530
;   LENGTH: 310
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004897.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  10
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
;   OTHER INFORMATION: NT HIT: AL163280.2, EVALUE 1.00e-105
;   OTHER INFORMATION: EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
;   OTHER INFORMATION: SWISSPROT HIT: P14077, EVALUE 1.00e-09
US-09-864-761-18530

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
              |||||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
              | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              |||||||||||||||||||||||| | ||||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              |||||||||||
Db        300 AGCAGGACTGA 310


RESULT 584
US-09-925-065A-829270/c
; Sequence 829270, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 829270
;   LENGTH: 309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-829270

  Query Match             3.4%;  Score 256.8;  DB 1;  Length 309;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  264;  Conservative    7;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:|||||||| |||||||||||| |||||||||||||||| | ||||||||
Db        309 AGAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCCACCTTTAAACACAGAGCTTGCAA 250

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||| |||||||||||| ||||||
Db        249 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATACTAATTAGGCAAAAACAGGA 190

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGG 7472
              ||||||||||||||||||:||||||:|||||||||||||||||||||| |||| ||
Db        189 AGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAAGGATCAGG 130

Qy       7473 ATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTA 7532
             |||||||||| |  ||||||| ||||| |||||||:|||||||||||||||| |||||
Db        129 ATATAAACCCAGGCATTCGAGCTGGCAAGGGCAACCCSCTTTGGGTCCCCTCCCCTTGTA 70

Qy       7533 TGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||  |||||||||||||||||||||:||||:
Db          69 TGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 20


RESULT 585
US-09-925-065A-829270/c
; Sequence 829270, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 829270
;   LENGTH: 309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-829270

  Query Match             3.4%;  Score 256.8;  DB 1;  Length 309;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  264;  Conservative    7;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
             |||||||:|||||||| |||||||||||| |||||||||||||||| | ||||||||
Db        309 AGAATCCCTAAGCCTAGTTGGGAAGGTGACCGCATCCACCTTTAAACACAGAGCTTGCAA 250

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:|||||||  |||||||||||||||||||||| |||||||||||| ||||||
Db        249 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATACTAATTAGGCAAAAACAGGA 190

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGG 7472
              ||||||||||||||||||:||||||:|||||||||||||||||||||| |||| ||
Db        189 AGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAAGGATCAGG 130

Qy       7473 ATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTA 7532
             |||||||||| |  ||||||| ||||| |||||||:|||||||||||||||| |||||
Db        129 ATATAAACCCAGGCATTCGAGCTGGCAAGGGCAACCCSCTTTGGGTCCCCTCCCCTTGTA 70

Qy       7533 TGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||  |||||||||||||||||||||:||||:
Db          69 TGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 20


RESULT 586
US-10-450-763-20002
; Sequence 20002, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 20002
;   LENGTH: 323
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (3)..(308)
;   OTHER INFORMATION: 72% homologous to multiple sclerosis associated retrovirus
;   OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=373.
US-10-450-763-20002

  Query Match             3.4%;  Score 251.2;  DB 1;  Length 323;
  Best Local Similarity   77.4%;  Pred. No. 0;
  Matches  278;  Conservative   23;  Mismatches   22;  Indels   36;  Gaps    1;

Qy       4008 GAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAG 4067
             ||||||||||:||||:|||||||||||||||:|||:| |||||| ||||||||||||||
Db          1 GAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCGGGCACCAGAATTAGGAGAAG 60

Qy       4068 RAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATG 4127
              :||||||||:|||||||||| || |||
Db         61 GAAAAAGGGTAAATATATATGCAAACT---------------------------------- 87

Qy       4128 MRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGG 4187
              ||||| |:|:|||||||||||||||||||||: ||:|||||| | |||:||||||||
Db         88 ---CAATAGGGAGAGAAAGGGAATTCCTAACTTCCAAGGGAACACTTGTCAAACATCAGG 144

Qy       4188 AAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACT 4247
             ||||||||| ||:|||||||:||||:||||||||||||:||||||||:||||||||||
Db        145 AAGCCATTAAGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACT 204

Qy       4248 GCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTG 4307
              ||:|||||||||| |||||||| :|||||||||||:|||:|||: ||||||| :||||||
Db        205 GCTGGGGTCATCAGAAAGGAAAAGAAAGGGAAATAGAAGAGAACCGCCAAGCGGATATTG 264

Qy       4308 AAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
             ||||:| |||||| ||||||||||||||||||||||||||||||    ||   ||| ||
Db        265 AAGCCAGAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGCTTACAGAAGGACCCCTA 323
```

```
RESULT 587
US-09-925-065A-585177
; Sequence 585177, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 585177
;   LENGTH: 279
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-585177

  Query Match             3.4%;  Score 251;  DB 1;  Length 279;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  257;  Conservative    6;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||| |||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db          1 CACTTGGGTTTTCCTGTTGAGAGCGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 60

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             | || |||||||||:|||||||||:|||||||||||  ||||||||||||||||||||||
Db         61 CGGAATAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTTAAACACGGGG 120

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 7405
             ||||||||||||:|||||||  |||||||||||||||||||||||||||||||||||||||
Db        121 CTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 180

Qy       7406 GACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAAT 7465
             |||||||||||||||||||||||||||:|||||||:|||||||||||||:||||||||||
Db        181 AACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCRGGAGGGACAAG 240

Qy       7466 GATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGC 7504
             |||| ||||||||||||| |  |||||||||||||| ||
Db        241 GATCAGGATATAAACCCAGGCATTCGAGCCGGCAACAGC 279


RESULT 588
US-09-925-065A-585177
; Sequence 585177, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 585177
;   LENGTH: 279
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-585177

  Query Match             3.4%;  Score 251;  DB 1;  Length 279;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  257;  Conservative    6;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||| |||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db          1 CACTTGGGTTTTCCTGTTGAGAGCGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 60

Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             | || |||||||||:|||||||||:|||||||||||  ||||||||||||||||||||||
Db         61 CGGAATAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTTAAACACGGGG 120

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 7405
             ||||||||||||:|||||||  |||||||||||||||||||||||||||||||||||||||
Db        121 CTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 180

Qy       7406 GACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAAT 7465
             |||||||||||||||||||||||||||:|||||||:|||||||||||||:||||||||||
Db        181 AACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCRGGAGGGACAAG 240

Qy       7466 GATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGC 7504
             |||| ||||||||||||| |  |||||||||||||| ||
Db        241 GATCAGGATATAAACCCAGGCATTCGAGCCGGCAACAGC 279


RESULT 589
US-10-029-386-17183
; Sequence 17183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR (
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17183
;   LENGTH: 333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR21_43.0
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.58
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.44
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.31
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 2.00e-10
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AA368311.1, EVALUE 1.00e-108
US-10-029-386-17183

  Query Match             3.3%;  Score 247.3;  DB 1;  Length 333;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  301;  Conservative    0;  Mismatches   32;  Indels    9;  Gaps    3;

Qy       2328 GAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACC 2387
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACC 60
```

```
Qy       2388 CGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGAC 2447
              ||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 TGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCCGAAGTCTGGGCAACAGAAGGAC 119

Qy       2448 AATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCC 2507
              |||| ||||||| |||||||||| || |||||||||||||||||||||||| | |||
Db        120 AATACGGACGAG-CAAAGAATGCCTGTGCTGTTCAAGTTAAACTAAAGGATTCCGCCTCC 178

Qy       2508 TTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATT 2567
              |||||||||||||||||||||||||| |||||  |||||||||||||| ||||||||||
Db        179 TTTCCCTACCAAAGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACTCCAAAAGATT 238

Qy       2568 GTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCC 2627
              ||||||||| |||||||||| ||| ||||||||||||||   |  ||||||| || | |
Db        239 GTTAAGGACCTAAAAGCCCATGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC- 297

Qy       2628 GTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAG 2669
                  |||    |  | ||||||||||   ||||||||||
Db        298 ------AATTTTAGGAGTACAGAAACCCAACAGACAGTGGAG 333


RESULT 590
US-09-864-761-30531
; Sequence 30531, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30531
;  LENGTH: 275
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: MAP TO AE000660.1
;  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;  OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99
;  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98
;  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
;  OTHER INFORMATION: EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
;  OTHER INFORMATION: NT HIT: AB024469.1, EVALUE 1.00e-118
;  OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 3.00e-17
US-09-864-761-30531

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              |||||||||||||||||:||:||||||| |||||:||||| ||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              |||||||| ||:||||:|||||:||||||||||: ||| |||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 591
US-09-925-065A-114790/c
; Sequence 114790, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114790
;  LENGTH: 317
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-925-065A-114790

  Query Match             3.2%;  Score 243.1;  DB 1;  Length 317;
  Best Local Similarity   79.3%;  Pred. No. 0;
```

```
  Matches  261;  Conservative   26;  Mismatches   25;  Indels   17;  Gaps    1;

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
              ||  | |    ||  | | || :  |||||||||||||||||:||||||||||||||||
Db        312 TTCCATGTCCCTGTACATCCTGACTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTT 253

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
              ||||||||:|||||||||||||||||||||||| |||||||||||||:||:|::||||||
Db        252 AGGGATAGCCCCCATCTATTTGGCCAGGCATTGGCCCAAGACTTGAGCCAGTTCTCATAC 193

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
              |||||||||||||||| || :|||  | |||||||||||||| :||:|||::|
Db        192 CTGGACACTCTTGTCTTTTGGTACATTGATGATTTACTTTAAGCCGCCCAT--------- 142

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
                    ||||||:||||||| :|||||:|||||||| ||:|||||  |||||||:||||
Db        141 --------TCAAGCYACCCAAGTGCTCTTAAATTTCCTTGCCACCTGATGCTACAAGGTT 90

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||||||||||||:||||:|||||||
Db         89 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 30

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYR 3178
              ||||||||||:||||||||||||||::
Db         29 AAAGGCACCAGGGCCCTCAGTGAGGAATG 1


RESULT 592
US-09-925-065A-114790/c
; Sequence 114790, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114790
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-114790

  Query Match             3.2%;  Score 243.1;  DB 1;  Length 317;
  Best Local Similarity   79.3%;  Pred. No. 0;
  Matches  261;  Conservative   26;  Mismatches   25;  Indels   17;  Gaps    1;

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
              ||  | |    ||  | | || :  |||||||||||||||||:||||||||||||||||
Db        312 TTCCATGTCCCTGTACATCCTGACTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTT 253

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
              ||||||||:|||||||||||||||||||||||| |||||||||||||:||:|::||||||
Db        252 AGGGATAGCCCCCATCTATTTGGCCAGGCATTGGCCCAAGACTTGAGCCAGTTCTCATAC 193

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
              |||||||||||||||| || :|||  | |||||||||||||| :||:|||::|
Db        192 CTGGACACTCTTGTCTTTTGGTACATTGATGATTTACTTTAAGCCGCCCAT--------- 142

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
                    ||||||:||||||| :|||||:|||||||| ||:|||||  |||||||:||||
Db        141 --------TCAAGCYACCCAAGTGCTCTTAAATTTCCTTGCCACCTGATGCTACAAGGTT 90

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||||||||||||:||||:|||||||
Db         89 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 30

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYR 3178
              ||||||||||:||||||||||||||::
Db         29 AAAGGCACCAGGGCCCTCAGTGAGGAATG 1


RESULT 593
US-10-040-916-50
; Sequence 50, Application US/10040916
; Publication No. US20020146769A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Evans, Cheryl
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
;                            ENCODING THEM
;        NUMBER OF SEQUENCES: 71
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,916
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/887,029
;             FILING DATE: 07-FEB-1997
;             APPLICATION NUMBER: 08/686,878
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 50:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 279 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-040-916-50

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
              :|||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60
```

```
Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
              :|||||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 594
US-08-979-847-100
; Sequence 100, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  100:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  310 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-100

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
              ||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||||

Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
              ||||||| ||| ||||| || |||||||| |||    | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
              ||||||||| ||||| |||||||||  |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| ||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 595
US-10-114-104-100
; Sequence 100, Application US/10114104
; Publication No. US20030198647A1
;  GENERAL INFORMATION:
;       APPLICANT: PERRON, HERVE
;                  BESEME, FREDERIC
;                  BEDIN, FREDERIC
;                  PARANHOS-BACCALA, GLAUCIA
;                  KOMURIAN-PRADEL, FLORENCE
;                  JOLIVET-REYNAUD, COLETTE
;                  MANDRAND, BERNARD
;                  GARSON, JEREMY
;                  TUKE, PHILIP
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL.
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 210
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OLIFF & BERRIDGE, PLC
;            STREET: P.O. BOX 19928
;            CITY: ALEXANDRIA
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22320
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/114,104
;            FILING DATE: 03-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/979,847
;            FILING DATE: 26-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BERRIDGE, WILLIAM P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 39046A
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;  INFORMATION FOR SEQ ID NO: 100:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 310 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
```

```
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 100:
US-10-114-104-100

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
              ||||||  ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
              ||||||| ||| ||||| || |||||||| |||    | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
              ||||||||| ||||| |||||||||||  |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| |||||||||||||:|||||||||.||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 596
US-10-029-386-14268/c
; Sequence 14268, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14268
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL137799.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.32
;   OTHER INFORMATION: EST_HUMAN HIT: AU121943.1, EVALUE 5.00e-93
;   OTHER INFORMATION: SWISSPROT HIT: P03355, EVALUE 6.00e-19
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 5.00e-95
US-10-029-386-14268

  Query Match             3.2%;  Score 236;  DB 1;  Length 266;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  233;  Conservative   19;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       2857 ATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATA 2916
              || |||| ||||||:| |||||||||||||||||:||||||||||||||| ||||||
Db        266 ATCCTTCGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCCGGGATA 207

Qy       2917 GYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACA 2976
              |: |||||||||||||||||| ||| ||||||||||||:||:|::|||||||| ||||

Db        206 GCTCCCATCTATTTGGCCAGGCGTTAACCCAAGACTTGAGCCAGTTCTCATACCTTGACA 147

Qy       2977 CTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCC 3036
              |||||||||||||:|||:|||||||||||| |||||:||: ||::|||||||||||||||||
Db        146 CTCTTGTCCTTCAGTATGTGGATGATTTGCTTTTAGCCACCCATTCAGAAACCTTGTGCC 87

Qy       3037 ATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAAC 3096
              ||||||| |||||||:||||| ||||||||||:|||||||||||||:||||||||||
Db         86 ATCAAGCTACCCAAGCGCTCTTGAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAAC 27

Qy       3097 SARARGCTCARCTCTGCTCACAGCAG 3122
              |:|:|||||:|||||||||| |||
Db         26 AAAAGGCTCAGCTCTGCTCACAACAG 1


RESULT 597
US-10-242-535A-31708
; Sequence 31708, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31708
;   LENGTH: 290
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-31708

  Query Match             3.1%;  Score 235.6;  DB 1;  Length 290;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  256;  Conservative    6;  Mismatches   15;  Indels   10;  Gaps    1;

Qy       7295 AATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACT 7354
              |||||:|||||||||:|||||||||||||| | |||||||||||||||||||||||||||
Db         14 AATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACT 73

Qy       7355 TAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGG 7414
              |||:||||||| |||||||||||||||||||||||||||||||| ||||| ||||||||
Db         74 TAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTAATTAAGCAAAAACAGGAGG 133

Qy       7415 TAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGAT 7474
              |||||||||||||| ||||:||| ||:|||||||||||| ||||||||||||||||||||
Db        134 TAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGCAAGAGGGACAATGATCGGGAT 193

Qy       7475 ATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATG 7534
              ||||||||| |  |||||||| |||||||| |||| ||||||||||||||||||||||||
Db        194 ATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATG 253

Qy       7535 GGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||| ||||          |||||||||||||||||||||:||||
Db        254 GGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 290


RESULT 598
US-10-085-783A-31708
; Sequence 31708, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
```

```
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31708
;   LENGTH: 290
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-31708

  Query Match             3.1%;  Score 235.6;  DB 1;  Length 290;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  256;  Conservative    6;  Mismatches   15;  Indels   10;  Gaps    1;

Qy       7295 AATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACT 7354
             |||||:|||||||||:|||||||||||||| | ||||||||||||||||||||||||||||
Db         14 AATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACT 73

Qy       7355 TAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGG 7414
             |||:||||||| ||||||||||||||||||||||||||||||||||| ||||| ||||||||
Db         74 TAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTAATTAAGCAAAAACAGGAGG 133

Qy       7415 TAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGAT 7474
             ||||||||||||||| ||||:||| ||:|||||||||||||| ||||||||||||||||||
Db        134 TAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGCAAGAGGGACAATGATCGGGAT 193

Qy       7475 ATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATG 7534
             ||||||||| |  |||||||| |||||||| |||| ||||||||||||||||||||||||
Db        194 ATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATG 253

Qy       7535 GGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||| ||||          ||||||||||||||||||:||||
Db        254 GGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 290


RESULT 599
US-08-979-847-2
; Sequence 2, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  |||||||| |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:|||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 600
US-10-430-442-2
; Sequence 2, Application US/10430442
; Publication No. US20030186391A1
;  GENERAL INFORMATION:
;       APPLICANT: Herve PERRON
;                  Frederic BESEME
;                  Frederic BEDIN
;                  Glaucia PARANHOS-BACCALA
;                  Florence KOMURIAN-PRADEL
;                  Colette JOLIVET
;                  Bernard MANDRAND
;       TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                           ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL.
;                           THERAPEUTIC PURPOSES
;       NUMBER OF SEQUENCES: 92
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Oliff & Berridge
;            STREET: 700 South Washington Street, Suite 300
;            CITY: Alexandria
;            STATE: Virginia
;            COUNTRY: U.S.A.
;            ZIP: 22314
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/430,442
;            FILING DATE: 07-May-2003
;          PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/691,563
;            FILING DATE: 02-AUG-1996
;          ATTORNEY/AGENT INFORMATION:
;            NAME: Berridge, William P.
;            REGISTRATION NUMBER: 30,024
;            REFERENCE/DOCKET NUMBER: WPB 38588
;          TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6400
;            TELEFAX: 703-836-2787
;     INFORMATION FOR SEQ ID NO: 2:
;          SEQUENCE CHARACTERISTICS:
;            LENGTH: 297 base pairs
;            TYPE: nucleotide
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;          MOLECULE TYPE: cDNA
;          SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-430-442-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 601
US-10-114-104-2
; Sequence 2, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;          ADDRESSEE: OLIFF & BERRIDGE, PLC
;          STREET: P.O. BOX 19928
;          CITY: ALEXANDRIA
;          STATE: VA
;          COUNTRY: USA
;            ZIP: 22320
;        COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/10/114,104
;          FILING DATE: 03-Apr-2002
;          CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: 08/979,847
;          FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;          NAME: BERRIDGE, WILLIAM P.
;          REGISTRATION NUMBER: 30,024
;          REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 703-836-6400
;          TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;          LENGTH: 297 base pairs
;          TYPE: nucleotide
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-114-104-2

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||    |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | ||||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 602
US-09-864-761-17758
; Sequence 17758, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
```

```
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17758
;   LENGTH: 278
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL008629.9
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.4
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 4.00e-79
;   OTHER INFORMATION: EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 1.60e-02
US-09-864-761-17758

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||| |||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | |||||||||||||||||||||||| |||| |

Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 603
US-10-741-600-65917
; Sequence 65917, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65917
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-65917

  Query Match             2.4%;  Score 181.8;  DB 1;  Length 201;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  186;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
             |||||:||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db          1 CTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 60

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGG 7472
             ||||||||||||||||||||:||||||:|||||||||||: |||||||||  ||||||
Db         61 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGYGGGAGGGACAAGCATCGGG 120

Qy       7473 ATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTA 7532
             |||||||||||  |||||||||||||||||||| ||||||||||||||||||||||||
Db        121 ATATAAACCCAGGCATTCGAGCCGGCAACGGCAACCACCTTTGGGTCCCCTCCCTTTGTA 180

Qy       7533 TGGGAGCTCTGTTTTCATGCT 7553
             |||||||||||  |||||  ||
Db        181 TGGGAGCTCTGCTTTCACTCT 201


RESULT 604
US-09-864-761-20979/c
; Sequence 20979, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20979
;   LENGTH: 200
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL034410.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;   OTHER INFORMATION: NT HIT: gi|1419444, EVALUE 5.00e-78
;   OTHER INFORMATION: EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
;   OTHER INFORMATION: SWISSPROT HIT: P03360, EVALUE 2.00e-06
US-09-864-761-20979

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  |||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 605
US-11-124-368A-7793/c
; Sequence 7793, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7793
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-7793

  Query Match             2.3%;  Score 172.2;  DB 1;  Length 201;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  180;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160
             |||| | |||||||||||||||||| |||||||||||| ||||||||||||  |||||
Db        201 CCCATGCTCTGATGTTAATGACATCGAAGGCACCCCTCCCGAGGAAATCTCAATTGCACG 142

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220
             ||| |||||||| ||||||||||||||||||||||||||||:||: | ||||||||||||
Db        141 ACCCCTACTACGTCCCAATTCAGCAGGAAGCAGTTAGAGCRGTCTTTGGCCAACCTCCCC 82

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280
             |||||||||| |||||||||||||||| | || ||| ||||| ||||||||||||||||
Db          81 AACAGCACTTGGGTTTTCCTGTTGAGAGGTGGTACTAAGAGATAGGACTAGCTGGATTTC 22

Qy       7281 CTAGGCTGAYTAAGAATCCYT 7301
             |||||||||:|||||||||:|
Db          21 CTAGGCTGACTAAGAATCCCT 1


RESULT 606
US-11-124-368A-7851/c
; Sequence 7851, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7851
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-7851

  Query Match             2.3%;  Score 171;  DB 1;  Length 201;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  180;  Conservative    3;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7034 TACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGC 7093
             | |||||||||||| ||| |||||||| ||||||||||||||  ||||||||||||||
Db        201 TTCAAATGGAGCCCCAGAAGCAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGC 142
```

```
Qy       7094 CTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAG 7153
              |||:||||||| | |||||||||||||||||| |||||||:|||||| ||||||||||||
Db        141 CTGCTAGCCCATGCTCTGATGTTAATGACATCGAAGGCACSCCTCCCGAGGAAATCTCAA 82

Qy       7154 CTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAA 7213
              ||||| ||| |||||||| ||||||||||||||||||||||||||||||: | ||||||
Db          81 TTGCACGACCCCTACTACGTCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCTTTGGCCAA 22

Qy       7214 CCTCCCCCAACAGCACTTAGGT 7234
              ||||||||||||||||| |||
Db          21 CCTCCCCAACAGCACTTGGGT 1


RESULT 607
US-10-741-600-46190/c
; Sequence 46190, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46190
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-46190

  Query Match             2.2%;  Score 164;  DB 1;  Length 201;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  176;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
              |||||||||| | |||||  |||||||||  |||||||||||||||||||||| |||||
Db        201 GGCTCTGGAATATGGAAAGCCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCTAGTGC 142

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
              |||||||||||| |||||||||||||||| | |||||:|||||| |||| || || |
Db        141 AGTCTACAAGGACGCTTTAAAAAAGATTGTCCGAATAGAARTAAGCCACCCCCTCATCTA 82

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
              |||||||||| ||||| |||||||||||||||||||| |||||||||| |||||| | |
Db          81 TGCCCCTTATGTCAAGAGAATCACTGGAAGGCCCACTGTCCCAGGGGACGAAGGTCCTCT 22

Qy       2000 GAGTCAGAAGCCACTAAC 2017
              ||||||||||||||||||
Db          21 GAGTCAGAAGCCACTAAC 4


RESULT 608
US-10-995-561-50894/c
; Sequence 50894, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50894
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-50894

  Query Match             2.2%;  Score 164;  DB 1;  Length 201;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches  176;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1820 GGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGC 1879
              |||||||||| | |||||  |||||||||  |||||||||||||||||||||| |||||
Db        201 GGCTCTGGAATATGGAAAGCCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCTAGTGC 142

Qy       1880 GGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCA 1939
              |||||||||||| |||||||||||||||| | |||||:|||||| |||| || || |
Db        141 AGTCTACAAGGACGCTTTAAAAAAGATTGTCCGAATAGAARTAAGCCACCCCCTCATCTA 82

Qy       1940 TGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTT 1999
              |||||||||| ||||| |||||||||||||||||||| |||||||||| |||||| | |
Db          81 TGCCCCTTATGTCAAGAGAATCACTGGAAGGCCCACTGTCCCAGGGGACGAAGGTCCTCT 22

Qy       2000 GAGTCAGAAGCCACTAAC 2017
              ||||||||||||||||||
Db          21 GAGTCAGAAGCCACTAAC 4


RESULT 609
US-10-995-561-50897/c
; Sequence 50897, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50897
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-50897

  Query Match             2.2%;  Score 161.9;  DB 1;  Length 201;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  182;  Conservative    1;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       1904 GATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCAC 1963
              |||||||| | |||||||||||| |||| || || |||||||||||| ||||| |||||||
Db        201 GATTGTCCGAATAGAAGTAAGCCACCCCCTCATCTATGCCCCTTATGTCAAGAGAATCAC 142

Qy       1964 TGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATG 2023
              |||||||||||||| |||||||||| |||||| | |||||:|||||||||||| ||||||
Db        141 TGGAAGGCCCACTGTCCCAGGGGACGAAGGTCCTCTGAGTSAGAAGCCACTAA-CAGATG 83

Qy       2024 ATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACA 2083
              |||||||||||||||||||||||||||  ||||||| |||| |||||| ||||||||||||
Db          82 ATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGTGCCAGCCCATGTCATCACCCTCACA 23

Qy       2084 GAGCCCTGGGTATGCTTGACC 2104
              |||||||||||||||||||||
Db          22 GAGCCCTGGGTATGCTTGACC 2


RESULT 610
US-10-995-561-50899/c
; Sequence 50899, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
```

```
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50899
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-50899

  Query Match             2.2%;  Score 161.3;  DB 1;  Length 201;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  182;  Conservative    1;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTT 1947
             ||||| |||||||||||||||| | |||||||||||| |||| || || |||||||||
Db        201 AGGACGCTTTAAAAAAGATTGTCCGAATAGAAGTAAGCCACCCCCTCATCTATGCCCCTT 142

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
             || ||||| |||||||||||||||||||| |||||||||: |||||| | |||||||||
Db        141 ATGTCAAGAGAATCACTGGAAGGCCCACTGTCCCAGGGGAYGAAGGTCCTCTGAGTCAGA 82

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCA 2067
             ||||||||| |||||||||||||||||||||||||||||||  ||||||| |||| ||||
Db         81 AGCCACTAA-CAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGTGCCAGCCCA 23

Qy       2068 TGCCATCACCCTCACAGAGCCC 2089
             || ||||||||||||||||||
Db         22 TGTCATCACCCTCACAGAGCCC 1


RESULT 611
US-09-864-761-32986
; Sequence 32986, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32986
;   LENGTH: 202
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007567.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
;   OTHER INFORMATION: NT HIT: AF156961.1, EVALUE 2.00e-89
;   OTHER INFORMATION: EST_HUMAN HIT: T96421.1, EVALUE 1.00e-45
;   OTHER INFORMATION: SWISSPROT HIT: Q02457, EVALUE 8.50e+00
US-09-864-761-32986

  Query Match             2.1%;  Score 156.6;  DB 1;  Length 202;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  183;  Conservative    5;  Mismatches   14;  Indels    2;  Gaps    2;

Qy        973 TAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCC 1032
              |||||||| ||||||| | |||||||||||||||||||||||||  ||||| ||||| |
Db          1 TAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCCGTTTATGCCCCCT-C 59

Qy       1033 CAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCA 1092
              |||||||||||||||| |||||||||||||||||||||||||||||| |:||||::|||| ||
Db         60 CAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCA 119

Qy       1093 GACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              ||||| |||||||| |||| |||| ||||| |||||||||||||:|||||||||||||||
Db        120 GACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAACCCTGATGGCTATATT 179

Qy       1153 GRTGTTTTACAAGGGTTAGGACAA 1176
              |:|| ||||||||||||||||||||
Db        180 GATG-TTTACAAGGGTTAGGACAA 202


RESULT 612
US-09-864-761-30401
; Sequence 30401, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30401
;   LENGTH: 207
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL121572.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EST_HUMAN HIT: AA368311.1, EVALUE 4.00e-63
;   OTHER INFORMATION: SWISSPROT HIT: P31792, EVALUE 7.00e-03
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 1.00e-65
US-09-864-761-30401

  Query Match             2.1%;  Score 156.2;  DB 1;  Length 207;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  188;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       2369 ACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAA 2428
             ||||| ||||||||||||  |||||||| || ||||||||||||||||||||||||||||
Db          1 ACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGGAATTAATCCTGAA 59

Qy       2429 GTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAA 2488
             |||||||||||||||||||||||||||||| | |||||||||| |||  |||||||||
Db         60 GTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATCCCATTCAAGTTAA 118

Qy       2489 ACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCA 2548
             |||| |||||||  | ||||||||||||||||||||||| |||| |||||| ||||| |
Db        119 ACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTAGACCCGAGGCCGA 178

Qy       2549 ACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             |||||| ||||||||||||||||||||||
Db        179 TCAAGGACTCCAAAAGATTGTTAAGGAC 206


RESULT 613
US-11-124-368A-7852/c
; Sequence 7852, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7852
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-7852

  Query Match             2.1%;  Score 154.7;  DB 1;  Length 201;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  174;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    1;

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||| ||||||||||||
Db        201 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTTA 142

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTA 7387
             | ||||||||||||||||||:|||||| ||||| ||         ||||||||||||
Db        141 AGCACGGGGCTTGCAACTTAGCTCACACCAGACCAGTCAGRTAGTAAAGAGAGCTCACTA 82

Qy       7388 AAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAG 7447
             |||||||||||||| || |||||||||||||||||||||||||:|||  | ||||||
Db         81 AAATGCTAATTAGGCTAAAACAGGAGGTAAAGAAATAGCCAATCATCTATCACGTGAGAG 22

Qy       7448 CACAGCAGGAGGGACAATGAT 7468
             |||||  | ||||||||||||
Db         21 CACAGGGGAAGGGACAATGAT 1


RESULT 614
US-10-741-600-46337/c
; Sequence 46337, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46337
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-46337

  Query Match             2.1%;  Score 154.3;  DB 1;  Length 201;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  175;  Conservative    3;  Mismatches   22;  Indels    1;  Gaps    1;

Qy        277 CTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTC 336
             ||||  |||||||||||| |||||||||||| |||||||  |||||||||  ||||  | |
Db        201 CTCCTGATTGGGCTAAAAGCTTGCCATTGTTCCTGCACACCTAAGTGCCCAGGTTCATCC 142

Qy        337 TAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCT 396
             ||||  |||||||||||| ||||||  ||||| |||||||:||||||||||||||||||
Db        141 TAATCCAGCTGAACACTAGTCACTGCATTCCACGGTTCTCWTCTGTGACCCACGGCTTCT 82

Qy        397 AATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSC 456
             ||||:| |||||||||| |||||||||||||| |||||||||||||||| ||||||:|
Db         81 AATAGAGCTATAACACTCACCACATGGCCCAAGGTTCCATTCCTTGGAATCTGTGAGGCC 22

Qy        457 AACGAACTCCAGGTCAGAGAA 477
             || |||| ||||||||||||||
Db         21 AA-GAACCCCAGGTCAGAGAA 2


RESULT 615
US-10-995-561-51113/c
; Sequence 51113, Application US/10995561
; Publication No. US20050272054A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51113
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-51113

  Query Match             2.1%;  Score 154.3;  DB 1;  Length 201;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  175;  Conservative    3;  Mismatches   22;  Indels    1;  Gaps    1;

Qy        277 CTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTC 336
              ||||  |||||||||||| |||||||||||| |||||||  |||||||||  ||||  | |
Db        201 CTCCTGATTGGGCTAAAAGCTTGCCATTGTTCCTGCACACCTAAGTGCCCAGGTTCATCC 142

Qy        337 TAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCT 396
              ||||  |||||||||||| ||||||  ||||| ||||||||:||||||||||||||||||
Db        141 TAATCCAGCTGAACACTAGTCACTGCATTCCACGGTTCTCWTCTGTGACCCACGGCTTCT 82

Qy        397 AATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSC 456
              ||||:| |||||||||| |||||||||||||| |||||||||||||||| ||||||:|
Db         81 AATAGAGCTATAACACTCACCACATGGCCCAAGGTTCCATTCCTTGGAATCTGTGAGGCC 22

Qy        457 AACGAACTCCAGGTCAGAGAA 477
              || |||| ||||||||||||||
Db         21 AA-GAACCCCAGGTCAGAGAA 2


RESULT 616
US-10-741-600-46338/c
; Sequence 46338, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46338
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-46338

  Query Match             2.0%;  Score 150.7;  DB 1;  Length 201;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  172;  Conservative    4;  Mismatches   24;  Indels    1;  Gaps    1;

Qy        298 TGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTC 357
              |||||||||  |||||||  |||||||||  ||||  |  |||||  |||||||||||| ||
Db        201 TGCCATTGTTCCTGCACACCTAAGTGCCCAGGTTCATCCTAATCCAGCTGAACACTAGTC 142

Qy        358 ACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACC 417
              ||||  |||||  ||||||||||||||||||||||||||||:||:| ||||||||||  |||
Db        141 ACTGCATTCCACGGTTCTCTTCTGTGACCCACGGCTTCTAMTAGAGCTATAACACTCACC 82

Qy        418 ACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAA 477
              |||||||||||| |||||||||||||||||| ||||||:||| |||| ||||||||||||||
Db         81 ACATGGCCCAAGGTTCCATTCCTTGGAATCTGTGAGGCCAA-GAACCCCAGGTCAGAGAA 23

Qy        478 TACGARGCTTGCCACCATCTT 498
              |  |:||||||  |||||||
Db         22 CAAAAGGCTTGCTGCCATCTT 2


RESULT 617
US-10-995-561-51115/c
; Sequence 51115, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51115
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-51115

  Query Match             2.0%;  Score 150.7;  DB 1;  Length 201;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  172;  Conservative    4;  Mismatches   24;  Indels    1;  Gaps    1;

Qy        298 TGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTC 357
              |||||||||  |||||||  |||||||||  ||||  |  |||||  |||||||||||| ||
Db        201 TGCCATTGTTCCTGCACACCTAAGTGCCCAGGTTCATCCTAATCCAGCTGAACACTAGTC 142

Qy        358 ACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACC 417
              ||||  |||||  ||||||||||||||||||||||||||||:||:| ||||||||||  |||
Db        141 ACTGCATTCCACGGTTCTCTTCTGTGACCCACGGCTTCTAMTAGAGCTATAACACTCACC 82

Qy        418 ACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAA 477
              |||||||||||| |||||||||||||||||| ||||||:||| |||| ||||||||||||||
Db         81 ACATGGCCCAAGGTTCCATTCCTTGGAATCTGTGAGGCCAA-GAACCCCAGGTCAGAGAA 23

Qy        478 TACGARGCTTGCCACCATCTT 498
              |  |:||||||  |||||||
Db         22 CAAAAGGCTTGCTGCCATCTT 2


RESULT 618
US-10-719-993-20008
; Sequence 20008, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20008
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-20008

  Query Match             2.0%;  Score 147.5;  DB 1;  Length 201;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  179;  Conservative    3;  Mismatches   15;  Indels   11;  Gaps    2;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              ||||||  |||||||||||||||||||||||||  || |||||  |||||||||||||  |||||
Db          1 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTCGTCCTAATCGAGCTGAACACTAGTCACT 60

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
```

```
              |||||||| ||||||||||  |||||||| ||||||||||          :|| | | ||
Db         61 GGGTTCCACGGTTCTCTTCCATGACCCACAGCTTCTAATA----------RCTCATCGCA 110

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||:||| |||| ||||||||||||| ||
Db        111 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACAC 169

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||:|||||||||||||||||||||| |||
Db        170 GAGGCTTGCCACCATCTTGGAAGCCGCC 197


RESULT 619
US-10-719-993-53980
; Sequence 53980, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53980
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-53980

  Query Match             2.0%;  Score 147.5;  DB 1;  Length 201;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  179;  Conservative    3;  Mismatches   15;  Indels   11;  Gaps    2;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| |||||||||||||||||||||| || ||||| |||||||||||| |||||
Db          1 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTCGTCCTAATCGAGCTGAACACTAGTCACT 60

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||| |||||||||  |||||||| ||||||||||          :|| | | ||
Db         61 GGGTTCCACGGTTCTCTTCCATGACCCACAGCTTCTAATA----------RCTCATCGCA 110

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||:||| |||| ||||||||||||| ||
Db        111 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACAC 169

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||:|||||||||||||||||||||| |||
Db        170 GAGGCTTGCCACCATCTTGGAAGCCGCC 197


RESULT 620
US-10-741-600-39842
; Sequence 39842, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39842
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-39842

  Query Match             2.0%;  Score 147.5;  DB 1;  Length 201;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  179;  Conservative    3;  Mismatches   15;  Indels   11;  Gaps    2;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| |||||||||||||||||||||| || ||||| |||||||||||| |||||
Db          1 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTCGTCCTAATCGAGCTGAACACTAGTCACT 60

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||| |||||||||  |||||||| ||||||||||          :|| | | ||
Db         61 GGGTTCCACGGTTCTCTTCCATGACCCACAGCTTCTAATA----------RCTCATCGCA 110

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||:||| |||| ||||||||||||| ||
Db        111 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACAC 169

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||:|||||||||||||||||||||| |||
Db        170 GAGGCTTGCCACCATCTTGGAAGCCGCC 197


RESULT 621
US-10-741-600-71390
; Sequence 71390, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71390
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-71390

  Query Match             2.0%;  Score 147.5;  DB 1;  Length 201;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  179;  Conservative    3;  Mismatches   15;  Indels   11;  Gaps    2;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| |||||||||||||||||||||| || ||||| |||||||||||| |||||
Db          1 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTCGTCCTAATCGAGCTGAACACTAGTCACT 60

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||| |||||||||  |||||||| ||||||||||          :|| | | ||
Db         61 GGGTTCCACGGTTCTCTTCCATGACCCACAGCTTCTAATA----------RCTCATCGCA 110

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||||||||||||||||||||||||||||:||| |||| ||||||||||||| ||
Db        111 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGAACAC 169

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||:|||||||||||||||||||||| |||
Db        170 GAGGCTTGCCACCATCTTGGAAGCCGCC 197


RESULT 622
US-10-029-386-18012/c
; Sequence 18012, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
```

```
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18012
;   LENGTH: 169
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC021619.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
;   OTHER INFORMATION: EST_HUMAN HIT: H01372.1, EVALUE 5.00e-61
;   OTHER INFORMATION: NT HIT: AF156963.1, EVALUE 3.00e-61
;   OTHER INFORMATION: SWISSPROT HIT: P08682, EVALUE 5.00e-01
US-10-029-386-18012

  Query Match             1.9%;  Score 145.2;  DB 1;  Length 169;
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  150;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       6879 AAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGG 6938
             ||||||  |||||||| ||||||||||:|||||||||||||||||||||||||||||||
Db        169 AAACACCAGACCCTGGGACCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGG 110

Qy       6939 ACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGT 6998
             |||||||||  |||||||| || |||||||||||||||||||||||||:|||||||| |
Db        109 ACCTCTAGCGACTATAATACTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTTT 50

Qy       6999 TAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              ||| |||||||||||||||| ||||||||:|| ||||||||||
Db         49 TAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTACAAATGG 6


RESULT 623
US-10-029-386-25984/c
; Sequence 25984, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25984
;   LENGTH: 194
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010749.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.88
;   OTHER INFORMATION: NT HIT: AB033233.1, EVALUE 2.00e-69
;   OTHER INFORMATION: EST_HUMAN HIT: BE144547.1, EVALUE 2.00e-69
;   OTHER INFORMATION: SWISSPROT HIT: P14010, EVALUE 2.50e+00
US-10-029-386-25984

  Query Match             1.9%;  Score 141.9;  DB 1;  Length 194;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  165;  Conservative    3;  Mismatches   26;  Indels    1;  Gaps    1;

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||| | ||||||| | |||||| || | |||||||||||||||||  |   ||
Db        194 CATTGTTCCTGCGCAGCTAAGTCCTTGGGTTCGTCCCAATTGAGCTGAACACTAGTTGCT 135

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||| |||||||||||||||| ||| |||||||||:| |||||||||| ||| ||
Db        134 GGGTTCCACGGTTCTCTTCTGTGATCCATGGCTTCTAATAGAGCTATAACACTCACCGCA 75

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              ||||||||| |||||||||||||||||||:||| |||| |||| |||| ||| |
Db         74 TGGCCCAAGGTTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGATCAGGGAACAA 16

Qy        481 GARGCTTGCCACCAT 495
              |:|||||  ||||
Db         15 AAGGCTTGCTGCCAT 1


RESULT 624
US-08-979-847-36
; Sequence 36, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLAC
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-36

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  |||||:| ||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              ||||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150
```

```
RESULT 625
US-10-114-104-36
; Sequence 36, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 36:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 150 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-10-114-104-36

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
             |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  |||||||:| |||||||||||||||||:|||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 626
US-10-029-386-17995
; Sequence 17995, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17995
;   LENGTH: 174
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR21_43.0
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: NT HIT: gi7656876, EVALUE 5.00e-51
;   OTHER INFORMATION: SWISSPROT HIT: Q9ZL29, EVALUE 8.60e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BF308839.1, EVALUE 7.00e-51
US-10-029-386-17995

  Query Match             1.8%;  Score 135.6;  DB 1;  Length 174;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  155;  Conservative    1;  Mismatches   16;  Indels    2;  Gaps    1;

Qy        101 TATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTT 160
              |||| |||||||||||||||||||||||||| ||||| |||| ||| ||||||||||||
Db          1 TATTCAATCTTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACAGCTTGAGCTGAGCTTT 60

Qy        161 TGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCC--GCTGACTCCCATC 218
              |||| || |||||||||||||||||||| ||| |||| |||||||  ||||||| |||||
Db         61 CGCTCGCCTTCCACCACTGCTGTTTGCCGCCATCGCAGACCTGCCGTGCTGACTTCCATC 120

Qy        219 CCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATT 272
              ||||| |||| |||||||||||||||||||| ||||||||||:||||||||
Db        121 CCTCTAGATCTGGCAGGGTGTCCGCTGTGCTCTTGATCCAGCGAGGCGCCCATT 174


RESULT 627
US-09-864-761-30590
; Sequence 30590, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30590
;   LENGTH: 152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022067.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: gi4759007, EVALUE 3.00e-54
;   OTHER INFORMATION: EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
;   OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 5.00e-10
US-09-864-761-30590

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             ||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:|||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
             |||||||||||| | | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 628
US-09-864-761-33021
; Sequence 33021, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33021
;   LENGTH: 133
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007567.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
;   OTHER INFORMATION: SWISSPROT HIT: P21414, EVALUE 1.00e-05
;   OTHER INFORMATION: EST_HUMAN HIT: AU121943.1, EVALUE 2.00e-52
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 2.00e-57
US-09-864-761-33021

  Query Match             1.6%;  Score 121.8;  DB 1;  Length 133;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  120;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2857 ATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATA 2916
              || |||| |||||||:| |||||||||||| |||||||:||||||||||||||||||||||
Db          1 ATCCTTCGAACCCAACGTCTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATA 60

Qy       2917 GYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACA 2976
              |:|||||||||||||||||| ||||||||||||||||||:||:|::|||||||||||||
Db         61 GCCCCCATCTATTTGGCCAGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACA 120

Qy       2977 CTCTTGTCCTTCR 2989
              |||||||||||||:
Db        121 CTCTTGTCCTTCA 133


RESULT 629
US-09-864-761-18806
; Sequence 18806, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18806
;   LENGTH: 155
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL121572.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
;   OTHER INFORMATION: EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 8.00e-42
;   OTHER INFORMATION: SWISSPROT HIT: P52490, EVALUE 6.60e+00
US-09-864-761-18806

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              || | |||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
              | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 630
US-09-864-761-26323/c
; Sequence 26323, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26323
;   LENGTH: 135
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL139245.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
```

```
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;   OTHER INFORMATION: SWISSPROT HIT: P18080, EVALUE 2.40e+00
;   OTHER INFORMATION: NT HIT: gi7656876, EVALUE 2.00e-33
;   OTHER INFORMATION: EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35
US-09-864-761-26323

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | |||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db          75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db          15 CTCAA 11


RESULT 631
US-10-310-914A-9963
; Sequence 9963, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9963
;   LENGTH: 130
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-9963

  Query Match             1.3%;  Score 98;  DB 1;  Length 130;
  Best Local Similarity   72.5%;  Pred. No. 0;
  Matches   87;  Conservative   20;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        447 CCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGG 506
              |||:||||:|||  | | |||||:||||||| ||||:||::||||||:|:::|||||  ||
Db          1 CCGUGAGGCCAAGAAUCCCCAGGUCAGAGAACACGAGGCUUGCCACCAUCUUGGAAGUGG 60

Qy        507 CCTGCTACCRTCTTGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCC 566
              ||:||  ||:: ::||||||:||  ||||||||:|::||||||:|:  ||||||||||||
Db         61 CCUGCCGCCAUUUUGGAAGUGGCCCACCACCAUCUUGGGAGCUCUCGGAGCAAGGACCCC 120


RESULT 632
US-09-864-761-23758/c
; Sequence 23758, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23758
;   LENGTH: 110
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL049870.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
;   OTHER INFORMATION: NT HIT: AF208161.1, EVALUE 6.00e-36
;   OTHER INFORMATION: SWISSPROT HIT: P03384, EVALUE 7.00e-04
;   OTHER INFORMATION: EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36
US-09-864-761-23758

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
              |||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db          50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 633
US-09-864-761-26878/c
; Sequence 26878, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
```

```
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26878
;   LENGTH: 103
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007244.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
;   OTHER INFORMATION: EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
;   OTHER INFORMATION: NT HIT: AF009668.1, EVALUE 1.00e-36
;   OTHER INFORMATION: SWISSPROT HIT: P10272, EVALUE 8.00e-05
US-09-864-761-26878

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
              ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 634
US-09-864-761-17757
; Sequence 17757, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17757
;   LENGTH: 107
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL008629.9
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
```

```
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 1.00e-27
;   OTHER INFORMATION: EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
;   OTHER INFORMATION: SWISSPROT HIT: O35083, EVALUE 2.40e+00
US-09-864-761-17757

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 635
US-10-029-386-19320/c
; Sequence 19320, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19320
;   LENGTH: 104
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF045450.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: NT HIT: AL163280.2, EVALUE 6.00e-52
;   OTHER INFORMATION: EST_HUMAN HIT: BF591249.1, EVALUE 1.00e-26
;   OTHER INFORMATION: SWISSPROT HIT: O95838, EVALUE 6.10e+00
US-10-029-386-19320

  Query Match             1.1%;  Score 85;  DB 1;  Length 104;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches   91;  Conservative    1;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        168 CGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGAT 227
              |||||||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||
Db        104 CGTCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGAT 45

Qy        228 CCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCAT 271
              || |||| || || |||||||||||||||||| ||:||||||||
Db         44 CCGGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCAT 1


RESULT 636
US-08-979-847-7
; Sequence 7, Application US/08979847
; Publication No. US20030039664A1
;  GENERAL INFORMATION:
;    APPLICANT:  PERRON, HERVE
;    APPLICANT:  BESEME, FREDERIC
;    APPLICANT:  BEDIN, FREDERIC
;    APPLICANT:  PARANHOS-BACCALA, GLAUCIA
;    APPLICANT:  KOMURIAN-PRADEL, FLORENCE
;    APPLICANT:  JOLIVET-REYNAUD, COLETTE
;    APPLICANT:  MANDRAND, BERNARD
;    APPLICANT:  GARSON, JEREMY
;    APPLICANT:  TUKE, PHILIP
;    TITLE OF INVENTION:  VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;    TITLE OF INVENTION:  ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
;    TITLE OF INVENTION:  THERAPEUTIC PURPOSES
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE, PLC
;      STREET:  P.O. BOX 19928
;      CITY:  ALEXANDRIA
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,847
;      FILING DATE:  26-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERRIDGE, WILLIAM P.
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 39046A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-979-847-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
              :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
              :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 637
US-10-430-442-7
; Sequence 7, Application US/10430442
; Publication No. US20030186391A1
;   GENERAL INFORMATION:
;        APPLICANT: Herve PERRON
;                   Frederic BESEME
;                   Frederic BEDIN
;                   Glaucia PARANHOS-BACCALA
;                   Florence KOMURIAN-PRADEL
;                   Colette JOLIVET
;                   Bernard MANDRAND
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 92
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Oliff & Berridge
;             STREET: 700 South Washington Street, Suite 300
;             CITY: Alexandria
;             STATE: Virginia
```

```
;             COUNTRY: U.S.A.
;             ZIP: 22314
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/430,442
;             FILING DATE: 07-May-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/691,563
;             FILING DATE: 02-AUG-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Berridge, William P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 38588
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 111 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-430-442-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 638
US-10-114-104-7
; Sequence 7, Application US/10114104
; Publication No. US20030198647A1
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,104
;             FILING DATE: 03-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,847
;             FILING DATE: 26-NOV-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;    INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 111 base pairs
;             TYPE: nucleotide
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-114-104-7

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


Search completed: August  7, 2006, 17:28:28
Job time : 2238 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2006, 08:22:50 ; Search time 3072 Seconds
(without alignments)
4.547 Million cell updates/sec

Title: US-10-717-580-11-COPY
Perfect score: 7489.6
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1198 seqs, 921189 residues

Total number of hits satisfying chosen parameters: 2392

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Genbank/EMBL

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1194 summaries

Database : gedb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2890.1 | 38.6 | 2930 | 1 | AX355872 | ACCESSION:AX355872 |
| 2 | 2890.1 | 38.6 | 2930 | 1 | AF208161 | ACCESSION:AF208161 |
| 3 | 2889.1 | 38.6 | 2946 | 1 | AR177269 | ACCESSION:AR177269 |
| 4 | 2889.1 | 38.6 | 2946 | 1 | BD086133 | ACCESSION:BD086133 |
| 5 | 2889.1 | 38.6 | 2946 | 1 | BD232464 | ACCESSION:BD232464 |
| 6 | 2706.61 | 36.1 | 2938 | 1 | BD196236 | ACCESSION:BD196236 |
| 7 | 2706.61 | 36.1 | 2938 | 1 | BD267478 | ACCESSION:BD267478 |
| 8 | 2706.61 | 36.1 | 2938 | 1 | AX000957 | ACCESSION:AX000957 |
| 9 | 2706.61 | 36.1 | 2938 | 1 | AX027471 | ACCESSION:AX027471 |
| 10 | 2706.61 | 36.1 | 2938 | 1 | AF072499 | ACCESSION:AF072499 |
| 11 | 2626.2 | 35.1 | 3372 | 1 | BD196243 | ACCESSION:BD196243 |
| 12 | 2626.2 | 35.1 | 3372 | 1 | BD267485 | ACCESSION:BD267485 |
| 13 | 2626.2 | 35.1 | 3372 | 1 | AX000964 | ACCESSION:AX000964 |
| 14 | 2626.2 | 35.1 | 3372 | 1 | AX027478 | ACCESSION:AX027478 |
| 15 | 2626.2 | 35.1 | 3372 | 1 | AF072502 | ACCESSION:AF072502 |
| 16 | 2551.4 | 34.1 | 2599 | 1 | BD221806 | ACCESSION:BD221806 |
| 17 | 2551.4 | 34.1 | 2599 | 1 | AR699755 | ACCESSION:AR699755 |
| 18 | 2551.4 | 34.1 | 2599 | 1 | AX007978 | ACCESSION:AX007978 |
| 19 | 2315.4 | 30.9 | 2784 | 1 | BD221809 | ACCESSION:BD221809 |
| 20 | 2315.4 | 30.9 | 2784 | 1 | AR699758 | ACCESSION:AR699758 |
| 21 | 2315.4 | 30.9 | 2784 | 1 | AX007981 | ACCESSION:AX007981 |
| 22 | 2258.9 | 30.2 | 2694 | 1 | AF520483S2 | ACCESSION:AF520484 |
| 23 | 2258.9 | 30.2 | 2694 | 1 | AF520485S2 | ACCESSION:AF520486 |
| 24 | 2258.9 | 30.2 | 2694 | 1 | AF520493S2 | ACCESSION:AF520494 |
| 25 | 2258.9 | 30.2 | 2694 | 1 | AF520495S2 | ACCESSION:AF520496 |
| 26 | 2258.9 | 30.2 | 2694 | 1 | AF520497S2 | ACCESSION:AF520498 |
| 27 | 2258.9 | 30.2 | 2694 | 1 | AF520499S2 | ACCESSION:AF520500 |
| 28 | 2258.9 | 30.2 | 2694 | 1 | AF520501S2 | ACCESSION:AF520502 |
| 29 | 2258.9 | 30.2 | 2694 | 1 | AF520503S2 | ACCESSION:AF520504 |
| 30 | 2258.9 | 30.2 | 2694 | 1 | AF520507S2 | ACCESSION:AF520508 |
| 31 | 2258.9 | 30.2 | 2694 | 1 | AF520513S2 | ACCESSION:AF520514 |
| 32 | 2258.9 | 30.2 | 2694 | 1 | AF520515S2 | ACCESSION:AF520516 |
| 33 | 2258.9 | 30.2 | 2694 | 1 | AF520525S2 | ACCESSION:AF520526 |
| 34 | 2258.9 | 30.2 | 2694 | 1 | AF520527S2 | ACCESSION:AF520528 |
| 35 | 2258.9 | 30.2 | 2694 | 1 | AF520531S2 | ACCESSION:AF520532 |
| 36 | 2258.9 | 30.2 | 2694 | 1 | AF520533S2 | ACCESSION:AF520534 |
| 37 | 2258.9 | 30.2 | 2694 | 1 | AF520537S2 | ACCESSION:AF520538 |
| 38 | 2258.9 | 30.2 | 2694 | 1 | AF520539S2 | ACCESSION:AF520540 |
| 39 | 2258.9 | 30.2 | 2694 | 1 | AF520541S2 | ACCESSION:AF520542 |
| 40 | 2258.9 | 30.2 | 2694 | 1 | AF520543S2 | ACCESSION:AF520544 |
| 41 | 2258.9 | 30.2 | 2694 | 1 | AF520547S2 | ACCESSION:AF520548 |
| 42 | 2258.9 | 30.2 | 2694 | 1 | AF520549S2 | ACCESSION:AF520550 |
| 43 | 2258.9 | 30.2 | 2694 | 1 | AF520551S2 | ACCESSION:AF520552 |
| 44 | 2257.3 | 30.1 | 2694 | 1 | AF520523S2 | ACCESSION:AF520524 |
| 45 | 2257.3 | 30.1 | 2694 | 1 | AF520535S2 | ACCESSION:AF520536 |
| 46 | 2257.3 | 30.1 | 2694 | 1 | AF520563S2 | ACCESSION:AF520564 |
| 47 | 2256.8 | 30.1 | 2372 | 1 | BD196244 | ACCESSION:BD196244 |
| 48 | 2256.8 | 30.1 | 2372 | 1 | BD267486 | ACCESSION:BD267486 |
| 49 | 2256.8 | 30.1 | 2372 | 1 | AX000965 | ACCESSION:AX000965 |
| 50 | 2256.8 | 30.1 | 2372 | 1 | AX027479 | ACCESSION:AX027479 |
| 51 | 2256.8 | 30.1 | 2372 | 1 | AF072504 | ACCESSION:AF072504 |
| 52 | 2255.7 | 30.1 | 2694 | 1 | AF520477S2 | ACCESSION:AF520478 |
| 53 | 2255.7 | 30.1 | 2694 | 1 | AF520479S2 | ACCESSION:AF520480 |
| 54 | 2255.7 | 30.1 | 2694 | 1 | AF520487S2 | ACCESSION:AF520488 |
| 55 | 2255.7 | 30.1 | 2694 | 1 | AF520505S2 | ACCESSION:AF520506 |
| 56 | 2255.7 | 30.1 | 2694 | 1 | AF520509S2 | ACCESSION:AF520510 |
| 57 | 2255.7 | 30.1 | 2694 | 1 | AF520511S2 | ACCESSION:AF520512 |
| 58 | 2255.7 | 30.1 | 2694 | 1 | AF520517S2 | ACCESSION:AF520518 |
| 59 | 2255.7 | 30.1 | 2694 | 1 | AF520519S2 | ACCESSION:AF520520 |
| 60 | 2255.7 | 30.1 | 2694 | 1 | AF520529S2 | ACCESSION:AF520530 |
| 61 | 2255.7 | 30.1 | 2694 | 1 | AF520553S2 | ACCESSION:AF520554 |
| 62 | 2255.7 | 30.1 | 2694 | 1 | AF520555S2 | ACCESSION:AF520556 |
| 63 | 2255.7 | 30.1 | 2694 | 1 | AF520557S2 | ACCESSION:AF520558 |
| 64 | 2255.7 | 30.1 | 2694 | 1 | AF520559S2 | ACCESSION:AF520560 |
| 65 | 2254.1 | 30.1 | 2694 | 1 | AF520489S2 | ACCESSION:AF520490 |
| 66 | 2254.1 | 30.1 | 2694 | 1 | AF520521S2 | ACCESSION:AF520522 |
| 67 | 2254.1 | 30.1 | 2694 | 1 | AF520561S2 | ACCESSION:AF520562 |
| 68 | 2252.5 | 30.1 | 2694 | 1 | AF520481S2 | ACCESSION:AF520482 |
| 69 | 2252.5 | 30.1 | 2694 | 1 | AF520545S2 | ACCESSION:AF520546 |
| 70 | 2250.9 | 30.1 | 2694 | 1 | AF520491S2 | ACCESSION:AF520492 |
| 71 | 2205.1 | 29.4 | 2782 | 1 | BD196241 | ACCESSION:BD196241 |
| 72 | 2205.1 | 29.4 | 2782 | 1 | BD267483 | ACCESSION:BD267483 |
| 73 | 2205.1 | 29.4 | 2782 | 1 | AX000962 | ACCESSION:AX000962 |
| 74 | 2205.1 | 29.4 | 2782 | 1 | AX027476 | ACCESSION:AX027476 |
| 75 | 2205.1 | 29.4 | 2782 | 1 | AX133396 | ACCESSION:AX133396 |
| 76 | 2201.9 | 29.4 | 2781 | 1 | AX092223 | ACCESSION:AX092223 |
| 77 | 2201.9 | 29.4 | 2781 | 1 | AF072506 | ACCESSION:AF072506 |
| 78 | 2141.4 | 28.6 | 2575 | 1 | BD196248 | ACCESSION:BD196248 |
| 79 | 2141.4 | 28.6 | 2575 | 1 | AX000969 | ACCESSION:AX000969 |
| 80 | 2109.4 | 28.2 | 2477 | 1 | AF123882 | ACCESSION:AF123882 |
| 81 | 2100.7 | 28.0 | 2778 | 1 | DQ256475 | ACCESSION:DQ256475 |
| 82 | 2080.3 | 27.8 | 2732 | 1 | DQ256474 | ACCESSION:DQ256474 |
| 83 | 2074.3 | 27.7 | 2764 | 1 | DQ256476 | ACCESSION:DQ256476 |
| 84 | 2036.1 | 27.2 | 2585 | 1 | BD196247 | ACCESSION:BD196247 |
| 85 | 2036.1 | 27.2 | 2585 | 1 | AX000968 | ACCESSION:AX000968 |
| 86 | 2032.5 | 27.1 | 2055 | 1 | BD221827 | ACCESSION:BD221827 |
| 87 | 2032.5 | 27.1 | 2055 | 1 | AR699776 | ACCESSION:AR699776 |
| 88 | 2032.5 | 27.1 | 2055 | 1 | AX007999 | ACCESSION:AX007999 |
| 89 | 1979.8 | 26.4 | 2346 | 1 | DQ256480 | ACCESSION:DQ256480 |
| 90 | 1977.8 | 26.4 | 2343 | 1 | DQ256481 | ACCESSION:DQ256481 |
| 91 | 1964.8 | 26.2 | 2341 | 1 | DQ256483 | ACCESSION:DQ256483 |
| 92 | 1957.1 | 26.1 | 2342 | 1 | DQ256482 | ACCESSION:DQ256482 |
| 93 | 1950.4 | 26.0 | 2344 | 1 | DQ256477 | ACCESSION:DQ256477 |
| 94 | 1906.3 | 25.5 | 2309 | 1 | DQ256479 | ACCESSION:DQ256479 |
| 95 | 1859.7 | 24.8 | 2319 | 1 | DQ256478 | ACCESSION:DQ256478 |
| 96 | 1859.6 | 24.8 | 2309 | 1 | AF135487 | ACCESSION:AF135487 |
| 97 | 1851.9 | 24.7 | 2304 | 1 | AR344375 | ACCESSION:AR344375 |
| 98 | 1851.9 | 24.7 | 2304 | 1 | AF009668 | ACCESSION:AF009668 |
| 99 | 1848.6 | 24.7 | 2006 | 1 | BD196238 | ACCESSION:BD196238 |
| 100 | 1848.6 | 24.7 | 2006 | 1 | BD267480 | ACCESSION:BD267480 |
| 101 | 1848.6 | 24.7 | 2006 | 1 | AX000959 | ACCESSION:AX000959 |
| 102 | 1848.6 | 24.7 | 2006 | 1 | AX027473 | ACCESSION:AX027473 |
| 103 | 1848.6 | 24.7 | 2006 | 1 | AF072503 | ACCESSION:AF072503 |
| 104 | 1834.3 | 24.5 | 1860 | 1 | AF513360 | ACCESSION:AF513360 |
| 105 | 1812.4 | 24.2 | 2074 | 1 | AX478550 | ACCESSION:AX478550 |
| 106 | 1768.9 | 23.6 | 2055 | 1 | CQ719381 | ACCESSION:CQ719381 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 107 | 1607 | 21.5 | 1617 | 1 | AF156963 | ACCESSION:AF156963 |
| | 108 | 1600.3 | 21.4 | 1624 | 1 | AF506835 | ACCESSION:AF506835 |
| | 109 | 1551.5 | 20.7 | 1894 | 1 | CQ053593 | ACCESSION:CQ053593 |
| | 110 | 1551.5 | 20.7 | 1894 | 1 | CQ068741 | ACCESSION:CQ068741 |
| | 111 | 1551.5 | 20.7 | 1894 | 1 | CQ095807 | ACCESSION:CQ095807 |
| | 112 | 1551.5 | 20.7 | 1894 | 1 | CQ134553 | ACCESSION:CQ134553 |
| | 113 | 1551.5 | 20.7 | 1894 | 1 | CQ173048 | ACCESSION:CQ173048 |
| | 114 | 1551.5 | 20.7 | 1894 | 1 | CQ202240 | ACCESSION:CQ202240 |
| | 115 | 1551.5 | 20.7 | 1894 | 1 | CQ217821 | ACCESSION:CQ217821 |
| | 116 | 1551.5 | 20.7 | 1894 | 1 | CQ256381 | ACCESSION:CQ256381 |
| | 117 | 1551.5 | 20.7 | 1894 | 1 | CQ293475 | ACCESSION:CQ293475 |
| | 118 | 1551.5 | 20.7 | 1894 | 1 | CQ330413 | ACCESSION:CQ330413 |
| c | 119 | 1487.1 | 19.9 | 1799 | 1 | BD221810 | ACCESSION:BD221810 |
| c | 120 | 1487.1 | 19.9 | 1799 | 1 | AR699759 | ACCESSION:AR699759 |
| c | 121 | 1487.1 | 19.9 | 1799 | 1 | AX007982 | ACCESSION:AX007982 |
| | 122 | 1440.3 | 19.2 | 1948 | 1 | BD196239 | ACCESSION:BD196239 |
| | 123 | 1440.3 | 19.2 | 1948 | 1 | BD267481 | ACCESSION:BD267481 |
| | 124 | 1440.3 | 19.2 | 1948 | 1 | AX000960 | ACCESSION:AX000960 |
| | 125 | 1440.3 | 19.2 | 1948 | 1 | AX027474 | ACCESSION:AX027474 |
| | 126 | 1440.3 | 19.2 | 1948 | 1 | AF072505 | ACCESSION:AF072505 |
| | 127 | 1434.3 | 19.2 | 1923 | 1 | AF156961 | ACCESSION:AF156961 |
| | 128 | 1431 | 19.1 | 1629 | 1 | BD252040 | ACCESSION:BD252040 |
| | 129 | 1431 | 19.1 | 1629 | 1 | AX036992 | ACCESSION:AX036992 |
| | 130 | 1431 | 19.1 | 1629 | 1 | AF331500 | ACCESSION:AF331500 |
| | 131 | 1329 | 17.7 | 1422 | 1 | BD196237 | ACCESSION:BD196237 |
| | 132 | 1329 | 17.7 | 1422 | 1 | BD267479 | ACCESSION:BD267479 |
| | 133 | 1329 | 17.7 | 1422 | 1 | AX000958 | ACCESSION:AX000958 |
| | 134 | 1329 | 17.7 | 1422 | 1 | AX027472 | ACCESSION:AX027472 |
| | 135 | 1329 | 17.7 | 1422 | 1 | AF072501 | ACCESSION:AF072501 |
| | 136 | 1324.2 | 17.7 | 1481 | 1 | BD136197 | ACCESSION:BD136197 |
| | 137 | 1324.2 | 17.7 | 1481 | 1 | CS248641 | ACCESSION:CS248641 |
| | 138 | 1324.2 | 17.7 | 1481 | 1 | AR344387 | ACCESSION:AR344387 |
| | 139 | 1324.2 | 17.7 | 1481 | 1 | AX001027 | ACCESSION:AX001027 |
| | 140 | 1320.4 | 17.6 | 1600 | 1 | AR344431 | ACCESSION:AR344431 |
| | 141 | 1309.8 | 17.5 | 1600 | 1 | AR344430 | ACCESSION:AR344430 |
| | 142 | 1291.9 | 17.2 | 1597 | 1 | AR344429 | ACCESSION:AR344429 |
| | 143 | 1270.4 | 17.0 | 1321 | 1 | BD196235 | ACCESSION:BD196235 |
| | 144 | 1270.4 | 17.0 | 1321 | 1 | BD267477 | ACCESSION:BD267477 |
| | 145 | 1270.4 | 17.0 | 1321 | 1 | AX000956 | ACCESSION:AX000956 |
| | 146 | 1270.4 | 17.0 | 1321 | 1 | AX027470 | ACCESSION:AX027470 |
| | 147 | 1270.4 | 17.0 | 1321 | 1 | AF072500 | ACCESSION:AF072500 |
| | 148 | 1229.2 | 16.4 | 1472 | 1 | CQ073417 | ACCESSION:CQ073417 |
| | 149 | 1229.2 | 16.4 | 1472 | 1 | CQ104281 | ACCESSION:CQ104281 |
| | 150 | 1229.2 | 16.4 | 1472 | 1 | CQ226221 | ACCESSION:CQ226221 |
| | 151 | 1229.2 | 16.4 | 1472 | 1 | CQ264339 | ACCESSION:CQ264339 |
| | 152 | 1229.2 | 16.4 | 1472 | 1 | CQ301436 | ACCESSION:CQ301436 |
| | 153 | 1201.4 | 16.0 | 1326 | 1 | BD221807 | ACCESSION:BD221807 |
| | 154 | 1201.4 | 16.0 | 1326 | 1 | AR699756 | ACCESSION:AR699756 |
| | 155 | 1201.4 | 16.0 | 1326 | 1 | AX007979 | ACCESSION:AX007979 |
| | 156 | 1117.7 | 14.9 | 1511 | 1 | BD136207 | ACCESSION:BD136207 |
| | 157 | 1117.7 | 14.9 | 1511 | 1 | BD267460 | ACCESSION:BD267460 |
| | 158 | 1117.7 | 14.9 | 1511 | 1 | CS248653 | ACCESSION:CS248653 |
| | 159 | 1117.7 | 14.9 | 1511 | 1 | AR344397 | ACCESSION:AR344397 |
| | 160 | 1117.7 | 14.9 | 1511 | 1 | AX001039 | ACCESSION:AX001039 |
| | 161 | 1117.7 | 14.9 | 1511 | 1 | AX027451 | ACCESSION:AX027451 |
| | 162 | 1117.7 | 14.9 | 1511 | 1 | AF123881 | ACCESSION:AF123881 |
| | 163 | 1106.8 | 14.8 | 1136 | 1 | BD196240 | ACCESSION:BD196240 |
| | 164 | 1106.8 | 14.8 | 1136 | 1 | BD267482 | ACCESSION:BD267482 |
| | 165 | 1106.8 | 14.8 | 1136 | 1 | AX000961 | ACCESSION:AX000961 |
| | 166 | 1106.8 | 14.8 | 1136 | 1 | AX027475 | ACCESSION:AX027475 |
| | 167 | 1106.8 | 14.8 | 1136 | 1 | AF072508 | ACCESSION:AF072508 |
| | 168 | 1054.1 | 14.1 | 1344 | 1 | CQ745026 | ACCESSION:CQ745026 |
| | 169 | 960.2 | 12.8 | 1080 | 1 | BD221828 | ACCESSION:BD221828 |
| | 170 | 960.2 | 12.8 | 1080 | 1 | AR699777 | ACCESSION:AR699777 |
| | 171 | 960.2 | 12.8 | 1080 | 1 | AX008005 | ACCESSION:AX008005 |
| | 172 | 952.8 | 12.7 | 1167 | 1 | A60173 | ACCESSION:A60173 |
| | 173 | 952.8 | 12.7 | 1167 | 1 | A79523 | ACCESSION:A79523 |
| | 174 | 952.8 | 12.7 | 1167 | 1 | AR094748 | ACCESSION:AR094748 |
| | 175 | 952.8 | 12.7 | 1167 | 1 | AR343195 | ACCESSION:AR343195 |
| | 176 | 952.8 | 12.7 | 1167 | 1 | AR344348 | ACCESSION:AR344348 |
| | 177 | 948 | 12.7 | 1117 | 1 | CQ082196 | ACCESSION:CQ082196 |
| | 178 | 948 | 12.7 | 1117 | 1 | CQ116881 | ACCESSION:CQ116881 |
| | 179 | 948 | 12.7 | 1117 | 1 | CQ238747 | ACCESSION:CQ238747 |
| | 180 | 948 | 12.7 | 1117 | 1 | CQ276425 | ACCESSION:CQ276425 |
| | 181 | 948 | 12.7 | 1117 | 1 | CQ313425 | ACCESSION:CQ313425 |
| | 182 | 936.9 | 12.5 | 1158 | 1 | A46012 | ACCESSION:A46012 |
| | 183 | 936.9 | 12.5 | 1158 | 1 | A57048 | ACCESSION:A57048 |
| | 184 | 936.9 | 12.5 | 1158 | 1 | A60113 | ACCESSION:A60113 |
| | 185 | 936.9 | 12.5 | 1158 | 1 | A79463 | ACCESSION:A79463 |
| | 186 | 936.9 | 12.5 | 1158 | 1 | A80197 | ACCESSION:A80197 |
| | 187 | 936.9 | 12.5 | 1158 | 1 | A80306 | ACCESSION:A80306 |
| | 188 | 936.9 | 12.5 | 1158 | 1 | AR035349 | ACCESSION:AR035349 |
| | 189 | 936.9 | 12.5 | 1158 | 1 | AR036190 | ACCESSION:AR036190 |
| | 190 | 936.9 | 12.5 | 1158 | 1 | AR036710 | ACCESSION:AR036710 |
| | 191 | 936.9 | 12.5 | 1158 | 1 | AR076972 | ACCESSION:AR076972 |
| | 192 | 936.9 | 12.5 | 1158 | 1 | AR094693 | ACCESSION:AR094693 |
| | 193 | 936.9 | 12.5 | 1158 | 1 | AR183928 | ACCESSION:AR183928 |
| | 194 | 936.9 | 12.5 | 1158 | 1 | AR343140 | ACCESSION:AR343140 |
| | 195 | 936.9 | 12.5 | 1158 | 1 | AR344297 | ACCESSION:AR344297 |
| | 196 | 914.6 | 12.2 | 957 | 1 | AY050301 | ACCESSION:AY050301 |
| | 197 | 913 | 12.2 | 957 | 1 | AY050299 | ACCESSION:AY050299 |
| | 198 | 912 | 12.2 | 957 | 1 | AY050300 | ACCESSION:AY050300 |
| | 199 | 897.601 | 12.0 | 1080 | 1 | CQ745034 | ACCESSION:CQ745034 |
| c | 200 | 894.401 | 11.9 | 1216 | 1 | BD221812 | ACCESSION:BD221812 |
| c | 201 | 894.401 | 11.9 | 1216 | 1 | AR699761 | ACCESSION:AR699761 |
| c | 202 | 894.401 | 11.9 | 1216 | 1 | AX007984 | ACCESSION:AX007984 |
| | 203 | 810.2 | 10.8 | 918 | 1 | CQ746663 | ACCESSION:CQ746663 |
| | 204 | 787.1 | 10.5 | 1038 | 1 | CQ747448 | ACCESSION:CQ747448 |
| | 205 | 747.7 | 10.0 | 918 | 1 | CQ746367 | ACCESSION:CQ746367 |
| | 206 | 741 | 9.9 | 993 | 1 | CQ749099 | ACCESSION:CQ749099 |
| | 207 | 740.4 | 9.9 | 892 | 1 | BV465687 | ACCESSION:BV465687 |
| | 208 | 715.301 | 9.6 | 870 | 1 | BV619299 | ACCESSION:BV619299 |
| | 209 | 712.9 | 9.5 | 792 | 1 | CQ058583 | ACCESSION:CQ058583 |
| | 210 | 712.9 | 9.5 | 792 | 1 | CQ077936 | ACCESSION:CQ077936 |
| | 211 | 712.9 | 9.5 | 792 | 1 | CQ108942 | ACCESSION:CQ108942 |
| | 212 | 712.9 | 9.5 | 792 | 1 | CQ147575 | ACCESSION:CQ147575 |
| | 213 | 712.9 | 9.5 | 792 | 1 | CQ182942 | ACCESSION:CQ182942 |
| | 214 | 712.9 | 9.5 | 792 | 1 | CQ207358 | ACCESSION:CQ207358 |
| | 215 | 712.9 | 9.5 | 792 | 1 | CQ230826 | ACCESSION:CQ230826 |
| | 216 | 712.9 | 9.5 | 792 | 1 | CQ268963 | ACCESSION:CQ268963 |
| | 217 | 712.9 | 9.5 | 792 | 1 | CQ305991 | ACCESSION:CQ305991 |
| | 218 | 712.9 | 9.5 | 792 | 1 | CQ343148 | ACCESSION:CQ343148 |
| | 219 | 699.7 | 9.3 | 795 | 1 | BV208491 | ACCESSION:BV208491 |
| | 220 | 689.4 | 9.2 | 772 | 1 | AB063618 | ACCESSION:AB063618 |
| | 221 | 674.5 | 9.0 | 944 | 1 | CQ073091 | ACCESSION:CQ073091 |
| | 222 | 674.5 | 9.0 | 944 | 1 | CQ103917 | ACCESSION:CQ103917 |
| | 223 | 674.5 | 9.0 | 944 | 1 | CQ225945 | ACCESSION:CQ225945 |
| | 224 | 674.5 | 9.0 | 944 | 1 | CQ301087 | ACCESSION:CQ301087 |
| | 225 | 660.7 | 8.8 | 787 | 1 | AB063621 | ACCESSION:AB063621 |
| c | 226 | 659.4 | 8.8 | 841 | 1 | BV469300 | ACCESSION:BV469300 |
| | 227 | 655 | 8.7 | 824 | 1 | BV457035 | ACCESSION:BV457035 |
| c | 228 | 653 | 8.7 | 749 | 1 | AY836220 | ACCESSION:AY836220 |
| | 229 | 643.5 | 8.6 | 808 | 1 | BV598358 | ACCESSION:BV598358 |
| c | 230 | 624.6 | 8.3 | 751 | 1 | AY836221 | ACCESSION:AY836221 |
| | 231 | 621.5 | 8.3 | 775 | 1 | BV552074 | ACCESSION:BV552074 |
| | 232 | 619.9 | 8.3 | 783 | 1 | AB063620 | ACCESSION:AB063620 |
| c | 233 | 617.3 | 8.2 | 759 | 1 | BV470413 | ACCESSION:BV470413 |
| | 234 | 613.5 | 8.2 | 765 | 1 | BV627795 | ACCESSION:BV627795 |
| | 235 | 608.1 | 8.1 | 711 | 1 | BV500243 | ACCESSION:BV500243 |
| c | 236 | 606.6 | 8.1 | 790 | 1 | BV667411 | ACCESSION:BV667411 |
| | 237 | 603.3 | 8.1 | 725 | 1 | BV487214 | ACCESSION:BV487214 |
| c | 238 | 598 | 8.0 | 771 | 1 | BV475768 | ACCESSION:BV475768 |
| c | 239 | 592.7 | 7.9 | 717 | 1 | AY836231 | ACCESSION:AY836231 |
| | 240 | 589.4 | 7.9 | 780 | 1 | AB063619 | ACCESSION:AB063619 |
| | 241 | 586 | 7.8 | 741 | 1 | A46020 | ACCESSION:A46020 |
| | 242 | 586 | 7.8 | 741 | 1 | A57056 | ACCESSION:A57056 |
| | 243 | 586 | 7.8 | 741 | 1 | A60121 | ACCESSION:A60121 |
| | 244 | 586 | 7.8 | 741 | 1 | A79471 | ACCESSION:A79471 |
| | 245 | 586 | 7.8 | 741 | 1 | A80205 | ACCESSION:A80205 |
| | 246 | 586 | 7.8 | 741 | 1 | A80314 | ACCESSION:A80314 |
| | 247 | 586 | 7.8 | 741 | 1 | AR035357 | ACCESSION:AR035357 |
| | 248 | 586 | 7.8 | 741 | 1 | AR036198 | ACCESSION:AR036198 |
| | 249 | 586 | 7.8 | 741 | 1 | AR036718 | ACCESSION:AR036718 |
| | 250 | 586 | 7.8 | 741 | 1 | AR076980 | ACCESSION:AR076980 |
| | 251 | 586 | 7.8 | 741 | 1 | AR094701 | ACCESSION:AR094701 |
| | 252 | 586 | 7.8 | 741 | 1 | AR183936 | ACCESSION:AR183936 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 253 | 586 | 7.8 | 741 | 1 | AR343148 | ACCESSION:AR343148 |
| | 254 | 586 | 7.8 | 741 | 1 | AR344305 | ACCESSION:AR344305 |
| | 255 | 585.501 | 7.8 | 778 | 1 | BV527275 | ACCESSION:BV527275 |
| | 256 | 579.9 | 7.7 | 661 | 1 | BV479791 | ACCESSION:BV479791 |
| c | 257 | 579.8 | 7.7 | 757 | 1 | BV470750 | ACCESSION:BV470750 |
| c | 258 | 578.9 | 7.7 | 705 | 1 | AY836232 | ACCESSION:AY836232 |
| | 259 | 577.9 | 7.7 | 787 | 1 | BV477069 | ACCESSION:BV477069 |
| | 260 | 575.9 | 7.7 | 792 | 1 | BV647045 | ACCESSION:BV647045 |
| | 261 | 572.8 | 7.6 | 781 | 1 | BV493485 | ACCESSION:BV493485 |
| c | 262 | 569.4 | 7.6 | 783 | 1 | BV650860 | ACCESSION:BV650860 |
| | 263 | 564.8 | 7.5 | 758 | 1 | BD136202 | ACCESSION:BD136202 |
| | 264 | 564.8 | 7.5 | 758 | 1 | CS248648 | ACCESSION:CS248648 |
| | 265 | 564.8 | 7.5 | 758 | 1 | AR344392 | ACCESSION:AR344392 |
| | 266 | 564.8 | 7.5 | 758 | 1 | AX001034 | ACCESSION:AX001034 |
| | 267 | 562.4 | 7.5 | 638 | 1 | AB115218 | ACCESSION:AB115218 |
| | 268 | 561.4 | 7.5 | 655 | 1 | AB115187 | ACCESSION:AB115187 |
| | 269 | 560.8 | 7.5 | 638 | 1 | AB115198 | ACCESSION:AB115198 |
| | 270 | 560.6 | 7.5 | 743 | 1 | BV607335 | ACCESSION:BV607335 |
| | 271 | 560.3 | 7.5 | 745 | 1 | BD146807 | ACCESSION:BD146807 |
| | 272 | 560.3 | 7.5 | 745 | 1 | AX866745 | ACCESSION:AX866745 |
| | 273 | 558.3 | 7.5 | 757 | 1 | BV600201 | ACCESSION:BV600201 |
| | 274 | 554.2 | 7.4 | 656 | 1 | AB115233 | ACCESSION:AB115233 |
| | 275 | 553.7 | 7.4 | 746 | 1 | BV595108 | ACCESSION:BV595108 |
| | 276 | 552 | 7.4 | 654 | 1 | AB056502 | ACCESSION:AB056502 |
| | 277 | 551.2 | 7.4 | 654 | 1 | AB115291 | ACCESSION:AB115291 |
| | 278 | 550.4 | 7.3 | 654 | 1 | AB115193 | ACCESSION:AB115193 |
| | 279 | 550.4 | 7.3 | 654 | 1 | AB115209 | ACCESSION:AB115209 |
| | 280 | 550.4 | 7.3 | 654 | 1 | AB115228 | ACCESSION:AB115228 |
| | 281 | 550.4 | 7.3 | 654 | 1 | AB115281 | ACCESSION:AB115281 |
| c | 282 | 549.1 | 7.3 | 716 | 1 | BV633838 | ACCESSION:BV633838 |
| | 283 | 548.8 | 7.3 | 654 | 1 | AB115249 | ACCESSION:AB115249 |
| | 284 | 548.4 | 7.3 | 638 | 1 | AB115186 | ACCESSION:AB115186 |
| | 285 | 548.4 | 7.3 | 640 | 1 | AB115197 | ACCESSION:AB115197 |
| | 286 | 547.8 | 7.3 | 654 | 1 | AB115214 | ACCESSION:AB115214 |
| | 287 | 547.6 | 7.3 | 654 | 1 | AB115290 | ACCESSION:AB115290 |
| | 288 | 547.2 | 7.3 | 654 | 1 | AB115215 | ACCESSION:AB115215 |
| | 289 | 547.2 | 7.3 | 654 | 1 | AB126104 | ACCESSION:AB126104 |
| | 290 | 545.6 | 7.3 | 654 | 1 | AB115254 | ACCESSION:AB115254 |
| | 291 | 543.4 | 7.3 | 655 | 1 | AB115194 | ACCESSION:AB115194 |
| | 292 | 542.4 | 7.2 | 654 | 1 | AB115210 | ACCESSION:AB115210 |
| | 293 | 541.7 | 7.2 | 657 | 1 | AB056495 | ACCESSION:AB056495 |
| | 294 | 541.7 | 7.2 | 657 | 1 | AB056498 | ACCESSION:AB056498 |
| | 295 | 541.7 | 7.2 | 657 | 1 | AB056500 | ACCESSION:AB056500 |
| | 296 | 541.3 | 7.2 | 640 | 1 | AB115220 | ACCESSION:AB115220 |
| | 297 | 540.1 | 7.2 | 657 | 1 | AB056493 | ACCESSION:AB056493 |
| | 298 | 540 | 7.2 | 656 | 1 | AB115229 | ACCESSION:AB115229 |
| | 299 | 539.9 | 7.2 | 655 | 1 | AB115280 | ACCESSION:AB115280 |
| | 300 | 538.4 | 7.2 | 540 | 1 | AB051004 | ACCESSION:AB051004 |
| | 301 | 538.4 | 7.2 | 540 | 1 | AB079501 | ACCESSION:AB079501 |
| | 302 | 538.4 | 7.2 | 540 | 1 | AB114919 | ACCESSION:AB114919 |
| | 303 | 538.4 | 7.2 | 540 | 1 | AB114943 | ACCESSION:AB114943 |
| | 304 | 538.4 | 7.2 | 540 | 1 | AB114952 | ACCESSION:AB114952 |
| | 305 | 538.4 | 7.2 | 540 | 1 | AB114961 | ACCESSION:AB114961 |
| | 306 | 538.4 | 7.2 | 540 | 1 | AB114970 | ACCESSION:AB114970 |
| | 307 | 538.4 | 7.2 | 540 | 1 | AB114973 | ACCESSION:AB114973 |
| | 308 | 538.4 | 7.2 | 540 | 1 | AB114989 | ACCESSION:AB114989 |
| | 309 | 537.6 | 7.2 | 654 | 1 | AB056490 | ACCESSION:AB056490 |
| | 310 | 537.5 | 7.2 | 643 | 1 | AB115196 | ACCESSION:AB115196 |
| | 311 | 537 | 7.2 | 650 | 1 | AB115204 | ACCESSION:AB115204 |
| | 312 | 536.9 | 7.2 | 657 | 1 | AB056492 | ACCESSION:AB056492 |
| | 313 | 536.8 | 7.2 | 540 | 1 | AB114922 | ACCESSION:AB114922 |
| | 314 | 536.8 | 7.2 | 540 | 1 | AB114937 | ACCESSION:AB114937 |
| | 315 | 536.8 | 7.2 | 540 | 1 | AB114945 | ACCESSION:AB114945 |
| | 316 | 536.8 | 7.2 | 540 | 1 | AB114948 | ACCESSION:AB114948 |
| | 317 | 536.8 | 7.2 | 540 | 1 | AB114959 | ACCESSION:AB114959 |
| | 318 | 536.8 | 7.2 | 540 | 1 | AB114966 | ACCESSION:AB114966 |
| | 319 | 536.8 | 7.2 | 540 | 1 | AB114968 | ACCESSION:AB114968 |
| | 320 | 536.8 | 7.2 | 540 | 1 | AB114978 | ACCESSION:AB114978 |
| | 321 | 536.8 | 7.2 | 540 | 1 | AB114979 | ACCESSION:AB114979 |
| | 322 | 536.5 | 7.2 | 656 | 1 | AB056496 | ACCESSION:AB056496 |
| | 323 | 536.5 | 7.2 | 656 | 1 | AB056501 | ACCESSION:AB056501 |
| | 324 | 536.4 | 7.2 | 650 | 1 | AB115203 | ACCESSION:AB115203 |
| | 325 | 536.4 | 7.2 | 650 | 1 | AB115260 | ACCESSION:AB115260 |
| | 326 | 535.6 | 7.2 | 544 | 1 | AB096025 | ACCESSION:AB096025 |
| | 327 | 535.2 | 7.1 | 540 | 1 | AB115000 | ACCESSION:AB115000 |
| | 328 | 534.9 | 7.1 | 656 | 1 | AB056503 | ACCESSION:AB056503 |
| | 329 | 534.6 | 7.1 | 656 | 1 | AB115287 | ACCESSION:AB115287 |
| | 330 | 534 | 7.1 | 542 | 1 | AB096022 | ACCESSION:AB096022 |
| | 331 | 534 | 7.1 | 654 | 1 | AB056486 | ACCESSION:AB056486 |
| | 332 | 533.6 | 7.1 | 540 | 1 | AB114947 | ACCESSION:AB114947 |
| | 333 | 533.4 | 7.1 | 648 | 1 | A57089 | ACCESSION:A57089 |
| | 334 | 533.4 | 7.1 | 648 | 1 | A57090 | ACCESSION:A57090 |
| | 335 | 533.4 | 7.1 | 648 | 1 | A80347 | ACCESSION:A80347 |
| | 336 | 533.4 | 7.1 | 648 | 1 | A80348 | ACCESSION:A80348 |
| | 337 | 533.3 | 7.1 | 640 | 1 | AB115219 | ACCESSION:AB115219 |
| | 338 | 532.9 | 7.1 | 656 | 1 | AB115279 | ACCESSION:AB115279 |
| | 339 | 532.1 | 7.1 | 649 | 1 | AB115212 | ACCESSION:AB115212 |
| | 340 | 532 | 7.1 | 540 | 1 | AB114930 | ACCESSION:AB114930 |
| | 341 | 531.7 | 7.1 | 653 | 1 | AB115300 | ACCESSION:AB115300 |
| | 342 | 531.6 | 7.1 | 650 | 1 | AB115232 | ACCESSION:AB115232 |
| | 343 | 531.6 | 7.1 | 650 | 1 | AB126119 | ACCESSION:AB126119 |
| | 344 | 531.2 | 7.1 | 642 | 1 | AB126113 | ACCESSION:AB126113 |
| | 345 | 530.9 | 7.1 | 649 | 1 | AB056484 | ACCESSION:AB056484 |
| | 346 | 530.6 | 7.1 | 650 | 1 | AB126115 | ACCESSION:AB126115 |
| | 347 | 530.5 | 7.1 | 656 | 1 | AB056497 | ACCESSION:AB056497 |
| | 348 | 530.4 | 7.1 | 540 | 1 | AB114991 | ACCESSION:AB114991 |
| | 349 | 530.1 | 7.1 | 656 | 1 | AB056494 | ACCESSION:AB056494 |
| | 350 | 530 | 7.1 | 650 | 1 | AB126112 | ACCESSION:AB126112 |
| | 351 | 530 | 7.1 | 650 | 1 | AB126114 | ACCESSION:AB126114 |
| | 352 | 530 | 7.1 | 650 | 1 | AB126116 | ACCESSION:AB126116 |
| c | 353 | 529.9 | 7.1 | 667 | 1 | BV491951 | ACCESSION:BV491951 |
| | 354 | 529.2 | 7.1 | 654 | 1 | AB056487 | ACCESSION:AB056487 |
| | 355 | 528.9 | 7.1 | 649 | 1 | AB056485 | ACCESSION:AB056485 |
| | 356 | 528.8 | 7.1 | 540 | 1 | AB114981 | ACCESSION:AB114981 |
| | 357 | 528.5 | 7.1 | 649 | 1 | AB115286 | ACCESSION:AB115286 |
| | 358 | 528.4 | 7.1 | 650 | 1 | AB126124 | ACCESSION:AB126124 |
| | 359 | 528 | 7.0 | 642 | 1 | AB126107 | ACCESSION:AB126107 |
| | 360 | 527.9 | 7.0 | 541 | 1 | AB079509 | ACCESSION:AB079509 |
| | 361 | 527.4 | 7.0 | 650 | 1 | AB126120 | ACCESSION:AB126120 |
| | 362 | 527.3 | 7.0 | 649 | 1 | AB056483 | ACCESSION:AB056483 |
| | 363 | 527.3 | 7.0 | 649 | 1 | AB115268 | ACCESSION:AB115268 |
| | 364 | 527.3 | 7.0 | 649 | 1 | AB115289 | ACCESSION:AB115289 |
| | 365 | 527.3 | 7.0 | 656 | 1 | AB056504 | ACCESSION:AB056504 |
| | 366 | 527.2 | 7.0 | 636 | 1 | AB115200 | ACCESSION:AB115200 |
| | 367 | 526.8 | 7.0 | 650 | 1 | AB126117 | ACCESSION:AB126117 |
| | 368 | 526.7 | 7.0 | 543 | 1 | AB096026 | ACCESSION:AB096026 |
| | 369 | 526.5 | 7.0 | 653 | 1 | AB126108 | ACCESSION:AB126108 |
| | 370 | 525.5 | 7.0 | 654 | 1 | AB056499 | ACCESSION:AB056499 |
| | 371 | 525.4 | 7.0 | 541 | 1 | AB115138 | ACCESSION:AB115138 |
| | 372 | 525.3 | 7.0 | 541 | 1 | AB114957 | ACCESSION:AB114957 |
| | 373 | 525.2 | 7.0 | 654 | 1 | AB056489 | ACCESSION:AB056489 |
| | 374 | 525.1 | 7.0 | 650 | 1 | AB115297 | ACCESSION:AB115297 |
| | 375 | 524.9 | 7.0 | 641 | 1 | AB115202 | ACCESSION:AB115202 |
| | 376 | 524.4 | 7.0 | 540 | 1 | AB114946 | ACCESSION:AB114946 |
| | 377 | 524.1 | 7.0 | 649 | 1 | AB115231 | ACCESSION:AB115231 |
| | 378 | 524 | 7.0 | 540 | 1 | AB114975 | ACCESSION:AB114975 |
| | 379 | 523.9 | 7.0 | 645 | 1 | A46019 | ACCESSION:A46019 |
| | 380 | 523.9 | 7.0 | 645 | 1 | A57055 | ACCESSION:A57055 |
| | 381 | 523.9 | 7.0 | 645 | 1 | A60120 | ACCESSION:A60120 |
| | 382 | 523.9 | 7.0 | 645 | 1 | A79470 | ACCESSION:A79470 |
| | 383 | 523.9 | 7.0 | 645 | 1 | A80204 | ACCESSION:A80204 |
| | 384 | 523.9 | 7.0 | 645 | 1 | A80313 | ACCESSION:A80313 |
| | 385 | 523.9 | 7.0 | 645 | 1 | AR035356 | ACCESSION:AR035356 |
| | 386 | 523.9 | 7.0 | 645 | 1 | AR036197 | ACCESSION:AR036197 |
| | 387 | 523.9 | 7.0 | 645 | 1 | AR036717 | ACCESSION:AR036717 |
| | 388 | 523.9 | 7.0 | 645 | 1 | AR076979 | ACCESSION:AR076979 |
| | 389 | 523.9 | 7.0 | 645 | 1 | AR094700 | ACCESSION:AR094700 |
| | 390 | 523.9 | 7.0 | 645 | 1 | AR183935 | ACCESSION:AR183935 |
| | 391 | 523.9 | 7.0 | 645 | 1 | AR343147 | ACCESSION:AR343147 |
| | 392 | 523.9 | 7.0 | 645 | 1 | AR344304 | ACCESSION:AR344304 |
| | 393 | 523.8 | 7.0 | 654 | 1 | AB115216 | ACCESSION:AB115216 |
| | 394 | 523.6 | 7.0 | 654 | 1 | AB056488 | ACCESSION:AB056488 |
| | 395 | 523.5 | 7.0 | 655 | 1 | AB115299 | ACCESSION:AB115299 |
| | 396 | 522.1 | 7.0 | 657 | 1 | AB115271 | ACCESSION:AB115271 |
| | 397 | 521.5 | 7.0 | 541 | 1 | AB079503 | ACCESSION:AB079503 |
| | 398 | 520.7 | 7.0 | 537 | 1 | AB115099 | ACCESSION:AB115099 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 399 | 520.7 | 7.0 | 537 | 1 | AB115169 | ACCESSION:AB115169 |
| 400 | 520 | 6.9 | 650 | 1 | AB115247 | ACCESSION:AB115247 |
| c 401 | 520 | 6.9 | 654 | 1 | BV485214 | ACCESSION:BV485214 |
| 402 | 519.9 | 6.9 | 541 | 1 | AB079504 | ACCESSION:AB079504 |
| 403 | 519.9 | 6.9 | 650 | 1 | AB115294 | ACCESSION:AB115294 |
| 404 | 519.6 | 6.9 | 650 | 1 | AB056491 | ACCESSION:AB056491 |
| 405 | 519.1 | 6.9 | 537 | 1 | AB054085 | ACCESSION:AB054085 |
| 406 | 519.1 | 6.9 | 537 | 1 | AB114963 | ACCESSION:AB114963 |
| 407 | 517.4 | 6.9 | 635 | 1 | BD136195 | ACCESSION:BD136195 |
| 408 | 517.4 | 6.9 | 635 | 1 | CS248638 | ACCESSION:CS248638 |
| 409 | 517.4 | 6.9 | 635 | 1 | AR344385 | ACCESSION:AR344385 |
| 410 | 517.4 | 6.9 | 635 | 1 | AX001024 | ACCESSION:AX001024 |
| 411 | 516.8 | 6.9 | 591 | 1 | BD196265 | ACCESSION:BD196265 |
| 412 | 516.8 | 6.9 | 591 | 1 | BD267476 | ACCESSION:BD267476 |
| 413 | 516.8 | 6.9 | 591 | 1 | AX000986 | ACCESSION:AX000986 |
| 414 | 516.8 | 6.9 | 591 | 1 | AX027469 | ACCESSION:AX027469 |
| 415 | 516.8 | 6.9 | 591 | 1 | AF072498 | ACCESSION:AF072498 |
| 416 | 516.8 | 6.9 | 650 | 1 | AB115292 | ACCESSION:AB115292 |
| 417 | 516 | 6.9 | 540 | 1 | AB114925 | ACCESSION:AB114925 |
| 418 | 515.8 | 6.9 | 650 | 1 | AB115266 | ACCESSION:AB115266 |
| 419 | 515.7 | 6.9 | 649 | 1 | AB115201 | ACCESSION:AB115201 |
| 420 | 515.3 | 6.9 | 652 | 1 | AB115270 | ACCESSION:AB115270 |
| 421 | 514.9 | 6.9 | 685 | 1 | AB115207 | ACCESSION:AB115207 |
| 422 | 514.4 | 6.9 | 540 | 1 | AB114953 | ACCESSION:AB114953 |
| 423 | 514.4 | 6.9 | 540 | 1 | AB115156 | ACCESSION:AB115156 |
| 424 | 514.2 | 6.9 | 650 | 1 | AB115269 | ACCESSION:AB115269 |
| 425 | 514.2 | 6.9 | 651 | 1 | AB115295 | ACCESSION:AB115295 |
| 426 | 513.2 | 6.9 | 638 | 1 | AB126118 | ACCESSION:AB126118 |
| 427 | 513.2 | 6.9 | 638 | 1 | AB126122 | ACCESSION:AB126122 |
| 428 | 513.2 | 6.9 | 638 | 1 | AB126127 | ACCESSION:AB126127 |
| 429 | 513.2 | 6.9 | 638 | 1 | AB126130 | ACCESSION:AB126130 |
| 430 | 512.9 | 6.8 | 539 | 1 | AB114934 | ACCESSION:AB114934 |
| 431 | 512.7 | 6.8 | 537 | 1 | AB054087 | ACCESSION:AB054087 |
| 432 | 512.1 | 6.8 | 651 | 1 | AB115206 | ACCESSION:AB115206 |
| 433 | 511.7 | 6.8 | 685 | 1 | AB115221 | ACCESSION:AB115221 |
| 434 | 511.7 | 6.8 | 685 | 1 | AB115304 | ACCESSION:AB115304 |
| 435 | 511.6 | 6.8 | 638 | 1 | AB126123 | ACCESSION:AB126123 |
| 436 | 511.6 | 6.8 | 638 | 1 | AB126126 | ACCESSION:AB126126 |
| 437 | 511.6 | 6.8 | 638 | 1 | AB126128 | ACCESSION:AB126128 |
| 438 | 511.6 | 6.8 | 638 | 1 | AB126129 | ACCESSION:AB126129 |
| 439 | 511.6 | 6.8 | 638 | 1 | AB126131 | ACCESSION:AB126131 |
| c 440 | 511.5 | 6.8 | 671 | 1 | BV486829 | ACCESSION:BV486829 |
| 441 | 511.2 | 6.8 | 540 | 1 | AB114928 | ACCESSION:AB114928 |
| 442 | 511.2 | 6.8 | 540 | 1 | AB114967 | ACCESSION:AB114967 |
| 443 | 511.1 | 6.8 | 537 | 1 | AB114923 | ACCESSION:AB114923 |
| 444 | 511.1 | 6.8 | 658 | 1 | AB115248 | ACCESSION:AB115248 |
| 445 | 510.7 | 6.8 | 658 | 1 | AB115285 | ACCESSION:AB115285 |
| 446 | 510.7 | 6.8 | 685 | 1 | AB115261 | ACCESSION:AB115261 |
| 447 | 510.6 | 6.8 | 668 | 1 | AB115199 | ACCESSION:AB115199 |
| 448 | 510.3 | 6.8 | 542 | 1 | BD003153 | ACCESSION:BD003153 |
| 449 | 510.3 | 6.8 | 542 | 1 | I80072 | ACCESSION:I80072 |
| 450 | 510.1 | 6.8 | 685 | 1 | AB115188 | ACCESSION:AB115188 |
| 451 | 510.1 | 6.8 | 685 | 1 | AB115192 | ACCESSION:AB115192 |
| 452 | 510.1 | 6.8 | 685 | 1 | AB115237 | ACCESSION:AB115237 |
| 453 | 510 | 6.8 | 638 | 1 | AB126101 | ACCESSION:AB126101 |
| 454 | 510 | 6.8 | 638 | 1 | AB126105 | ACCESSION:AB126105 |
| 455 | 510 | 6.8 | 638 | 1 | AB126106 | ACCESSION:AB126106 |
| 456 | 510 | 6.8 | 638 | 1 | AB126109 | ACCESSION:AB126109 |
| 457 | 510 | 6.8 | 638 | 1 | AB126111 | ACCESSION:AB126111 |
| 458 | 509.6 | 6.8 | 540 | 1 | AB115097 | ACCESSION:AB115097 |
| 459 | 509.1 | 6.8 | 685 | 1 | AB115238 | ACCESSION:AB115238 |
| 460 | 508.4 | 6.8 | 540 | 1 | AB115113 | ACCESSION:AB115113 |
| 461 | 508.4 | 6.8 | 638 | 1 | AB126110 | ACCESSION:AB126110 |
| 462 | 508.4 | 6.8 | 638 | 1 | AB126121 | ACCESSION:AB126121 |
| 463 | 508.4 | 6.8 | 642 | 1 | AB056482 | ACCESSION:AB056482 |
| 464 | 507.9 | 6.8 | 686 | 1 | AB115305 | ACCESSION:AB115305 |
| 465 | 507.7 | 6.8 | 539 | 1 | AB114931 | ACCESSION:AB114931 |
| 466 | 507.7 | 6.8 | 539 | 1 | AB114960 | ACCESSION:AB114960 |
| 467 | 507.4 | 6.8 | 639 | 1 | AB126102 | ACCESSION:AB126102 |
| 468 | 507.2 | 6.8 | 543 | 1 | AB054089 | ACCESSION:AB054089 |
| 469 | 507.2 | 6.8 | 543 | 1 | AB055416 | ACCESSION:AB055416 |
| 470 | 506.9 | 6.8 | 651 | 1 | AB115296 | ACCESSION:AB115296 |
| 471 | 506.9 | 6.8 | 685 | 1 | AB115236 | ACCESSION:AB115236 |
| 472 | 506.8 | 6.8 | 540 | 1 | AB115104 | ACCESSION:AB115104 |
| 473 | 506.8 | 6.8 | 642 | 1 | AB056481 | ACCESSION:AB056481 |
| 474 | 506.5 | 6.8 | 685 | 1 | AB115230 | ACCESSION:AB115230 |
| 475 | 506.4 | 6.8 | 540 | 1 | AB114969 | ACCESSION:AB114969 |
| 476 | 506.4 | 6.8 | 540 | 1 | AB115136 | ACCESSION:AB115136 |
| 477 | 506.4 | 6.8 | 654 | 1 | AB115211 | ACCESSION:AB115211 |
| 478 | 505.9 | 6.8 | 520 | 1 | AB120045 | ACCESSION:AB120045 |
| 479 | 505.9 | 6.8 | 520 | 1 | AB120051 | ACCESSION:AB120051 |
| 480 | 505.9 | 6.8 | 520 | 1 | AB120054 | ACCESSION:AB120054 |
| 481 | 505.9 | 6.8 | 520 | 1 | AB120056 | ACCESSION:AB120056 |
| 482 | 505.2 | 6.7 | 678 | 1 | BD196263 | ACCESSION:BD196263 |
| 483 | 505.2 | 6.7 | 678 | 1 | BD267475 | ACCESSION:BD267475 |
| 484 | 505.2 | 6.7 | 678 | 1 | AX000984 | ACCESSION:AX000984 |
| 485 | 505.2 | 6.7 | 678 | 1 | AX027468 | ACCESSION:AX027468 |
| 486 | 505.2 | 6.7 | 678 | 1 | AF072494 | ACCESSION:AF072494 |
| 487 | 505.1 | 6.7 | 537 | 1 | AB114918 | ACCESSION:AB114918 |
| 488 | 504.8 | 6.7 | 540 | 1 | AB051010 | ACCESSION:AB051010 |
| 489 | 504.8 | 6.7 | 540 | 1 | AB115112 | ACCESSION:AB115112 |
| 490 | 504.8 | 6.7 | 540 | 1 | AB115120 | ACCESSION:AB115120 |
| 491 | 504.3 | 6.7 | 520 | 1 | AB120044 | ACCESSION:AB120044 |
| 492 | 504.3 | 6.7 | 520 | 1 | AB120048 | ACCESSION:AB120048 |
| 493 | 504.3 | 6.7 | 520 | 1 | AB120049 | ACCESSION:AB120049 |
| 494 | 504.3 | 6.7 | 520 | 1 | AB120057 | ACCESSION:AB120057 |
| 495 | 503.5 | 6.7 | 537 | 1 | AB115116 | ACCESSION:AB115116 |
| 496 | 503.5 | 6.7 | 537 | 1 | AB115124 | ACCESSION:AB115124 |
| 497 | 503.2 | 6.7 | 540 | 1 | AB115103 | ACCESSION:AB115103 |
| 498 | 503.2 | 6.7 | 540 | 1 | AB115132 | ACCESSION:AB115132 |
| 499 | 503.2 | 6.7 | 540 | 1 | AB115150 | ACCESSION:AB115150 |
| 500 | 503.1 | 6.7 | 537 | 1 | AB115133 | ACCESSION:AB115133 |
| 501 | 503.1 | 6.7 | 686 | 1 | AB115303 | ACCESSION:AB115303 |
| 502 | 503 | 6.7 | 542 | 1 | AB079502 | ACCESSION:AB079502 |
| 503 | 503 | 6.7 | 684 | 1 | AB115213 | ACCESSION:AB115213 |
| 504 | 502.7 | 6.7 | 520 | 1 | AB120050 | ACCESSION:AB120050 |
| 505 | 502.7 | 6.7 | 520 | 1 | AB120053 | ACCESSION:AB120053 |
| 506 | 502.7 | 6.7 | 520 | 1 | AB120055 | ACCESSION:AB120055 |
| 507 | 502.4 | 6.7 | 685 | 1 | AB115224 | ACCESSION:AB115224 |
| 508 | 502.2 | 6.7 | 660 | 1 | AB126103 | ACCESSION:AB126103 |
| 509 | 502.1 | 6.7 | 652 | 1 | AB115284 | ACCESSION:AB115284 |
| 510 | 502 | 6.7 | 540 | 1 | AB055412 | ACCESSION:AB055412 |
| 511 | 502 | 6.7 | 540 | 1 | AB115121 | ACCESSION:AB115121 |
| 512 | 501.6 | 6.7 | 540 | 1 | AB115119 | ACCESSION:AB115119 |
| 513 | 501.6 | 6.7 | 540 | 1 | AB115122 | ACCESSION:AB115122 |
| 514 | 501.1 | 6.7 | 520 | 1 | AB120046 | ACCESSION:AB120046 |
| 515 | 500.4 | 6.7 | 540 | 1 | AB115144 | ACCESSION:AB115144 |
| 516 | 500.4 | 6.7 | 540 | 1 | AB115168 | ACCESSION:AB115168 |
| 517 | 500.1 | 6.7 | 640 | 1 | AB126125 | ACCESSION:AB126125 |
| 518 | 500 | 6.7 | 540 | 1 | AB115006 | ACCESSION:AB115006 |
| 519 | 500 | 6.7 | 686 | 1 | AB115306 | ACCESSION:AB115306 |
| 520 | 499.5 | 6.7 | 645 | 1 | AB115301 | ACCESSION:AB115301 |
| 521 | 499.1 | 6.7 | 539 | 1 | AB114987 | ACCESSION:AB114987 |
| 522 | 498.8 | 6.7 | 540 | 1 | AB115102 | ACCESSION:AB115102 |
| 523 | 498.5 | 6.7 | 539 | 1 | AB115007 | ACCESSION:AB115007 |
| 524 | 498.4 | 6.7 | 540 | 1 | AB051003 | ACCESSION:AB051003 |
| 525 | 498.4 | 6.7 | 540 | 1 | AB114962 | ACCESSION:AB114962 |
| 526 | 498.4 | 6.7 | 540 | 1 | AB114990 | ACCESSION:AB114990 |
| 527 | 498.4 | 6.7 | 540 | 1 | AB115130 | ACCESSION:AB115130 |
| 528 | 498.4 | 6.7 | 540 | 1 | AB115139 | ACCESSION:AB115139 |
| 529 | 498.4 | 6.7 | 540 | 1 | AB115143 | ACCESSION:AB115143 |
| 530 | 498.3 | 6.7 | 537 | 1 | AB114976 | ACCESSION:AB114976 |
| 531 | 498.3 | 6.7 | 537 | 1 | AB115153 | ACCESSION:AB115153 |
| 532 | 498.1 | 6.7 | 539 | 1 | AB114916 | ACCESSION:AB114916 |
| 533 | 497.9 | 6.6 | 520 | 1 | AB120047 | ACCESSION:AB120047 |
| 534 | 497.3 | 6.6 | 649 | 1 | AB115191 | ACCESSION:AB115191 |
| 535 | 497.2 | 6.6 | 540 | 1 | AB115126 | ACCESSION:AB115126 |
| 536 | 497.1 | 6.6 | 627 | 1 | CQ742266 | ACCESSION:CQ742266 |
| 537 | 496.8 | 6.6 | 540 | 1 | AB050998 | ACCESSION:AB050998 |
| 538 | 496.8 | 6.6 | 540 | 1 | AB051007 | ACCESSION:AB051007 |
| 539 | 496.8 | 6.6 | 540 | 1 | AB114954 | ACCESSION:AB114954 |
| 540 | 496.8 | 6.6 | 540 | 1 | AB115001 | ACCESSION:AB115001 |
| 541 | 496.8 | 6.6 | 540 | 1 | AB115117 | ACCESSION:AB115117 |
| 542 | 496.8 | 6.6 | 540 | 1 | AB115123 | ACCESSION:AB115123 |
| 543 | 496.8 | 6.6 | 540 | 1 | AB115142 | ACCESSION:AB115142 |
| 544 | 495.6 | 6.6 | 540 | 1 | AB115160 | ACCESSION:AB115160 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 545 | 495.2 | 6.6 | 540 | 1 | AB114982 | ACCESSION:AB114982 |
| 546 | 495.2 | 6.6 | 540 | 1 | AB114996 | ACCESSION:AB114996 |
| 547 | 495.2 | 6.6 | 540 | 1 | AB115141 | ACCESSION:AB115141 |
| 548 | 495.1 | 6.6 | 652 | 1 | AB115288 | ACCESSION:AB115288 |
| 549 | 494.7 | 6.6 | 520 | 1 | AB120052 | ACCESSION:AB120052 |
| 550 | 493.9 | 6.6 | 537 | 1 | AB115110 | ACCESSION:AB115110 |
| 551 | 493.9 | 6.6 | 537 | 1 | AB115154 | ACCESSION:AB115154 |
| 552 | 493.7 | 6.6 | 539 | 1 | AB114921 | ACCESSION:AB114921 |
| 553 | 493.6 | 6.6 | 540 | 1 | AB114999 | ACCESSION:AB114999 |
| 554 | 493.6 | 6.6 | 540 | 1 | AB115008 | ACCESSION:AB115008 |
| 555 | 493.6 | 6.6 | 540 | 1 | AB115009 | ACCESSION:AB115009 |
| 556 | 493.5 | 6.6 | 537 | 1 | AB115134 | ACCESSION:AB115134 |
| 557 | 493.3 | 6.6 | 539 | 1 | AB114926 | ACCESSION:AB114926 |
| 558 | 492.2 | 6.6 | 521 | 1 | AB120058 | ACCESSION:AB120058 |
| 559 | 492.1 | 6.6 | 539 | 1 | AB114939 | ACCESSION:AB114939 |
| 560 | 492 | 6.6 | 536 | 1 | AB114944 | ACCESSION:AB114944 |
| 561 | 492 | 6.6 | 540 | 1 | AB114956 | ACCESSION:AB114956 |
| 562 | 491.7 | 6.6 | 539 | 1 | AB114972 | ACCESSION:AB114972 |
| 563 | 491.7 | 6.6 | 539 | 1 | AB126149 | ACCESSION:AB126149 |
| 564 | 490.4 | 6.5 | 540 | 1 | AB115173 | ACCESSION:AB115173 |
| 565 | 490.3 | 6.5 | 537 | 1 | AB115170 | ACCESSION:AB115170 |
| 566 | 490.1 | 6.5 | 539 | 1 | AB115003 | ACCESSION:AB115003 |
| 567 | 489.9 | 6.5 | 541 | 1 | AB114942 | ACCESSION:AB114942 |
| 568 | 489.1 | 6.5 | 537 | 1 | AB115111 | ACCESSION:AB115111 |
| 569 | 489 | 6.5 | 550 | 1 | AB096023 | ACCESSION:AB096023 |
| 570 | 488.9 | 6.5 | 535 | 1 | AB114924 | ACCESSION:AB114924 |
| 571 | 488.9 | 6.5 | 539 | 1 | AB114935 | ACCESSION:AB114935 |
| 572 | 488.9 | 6.5 | 539 | 1 | AB114951 | ACCESSION:AB114951 |
| 573 | 488.9 | 6.5 | 539 | 1 | AB114977 | ACCESSION:AB114977 |
| 574 | 488.8 | 6.5 | 660 | 1 | AB115282 | ACCESSION:AB115282 |
| 575 | 488.5 | 6.5 | 539 | 1 | AB126160 | ACCESSION:AB126160 |
| 576 | 487.8 | 6.5 | 541 | 1 | AB115128 | ACCESSION:AB115128 |
| 577 | 487.5 | 6.5 | 537 | 1 | AB115118 | ACCESSION:AB115118 |
| 578 | 487.5 | 6.5 | 537 | 1 | AB115172 | ACCESSION:AB115172 |
| 579 | 487.3 | 6.5 | 539 | 1 | AB114917 | ACCESSION:AB114917 |
| 580 | 487.3 | 6.5 | 539 | 1 | AB114949 | ACCESSION:AB114949 |
| 581 | 487.3 | 6.5 | 539 | 1 | AB115151 | ACCESSION:AB115151 |
| 582 | 487.2 | 6.5 | 540 | 1 | AB051009 | ACCESSION:AB051009 |
| 583 | 487.2 | 6.5 | 540 | 1 | AB114920 | ACCESSION:AB114920 |
| 584 | 487.1 | 6.5 | 544 | 1 | AF127227 | ACCESSION:AF127227 |
| 585 | 486.9 | 6.5 | 539 | 1 | AB126142 | ACCESSION:AB126142 |
| 586 | 486.1 | 6.5 | 537 | 1 | AB126137 | ACCESSION:AB126137 |
| 587 | 485.9 | 6.5 | 537 | 1 | AB051005 | ACCESSION:AB051005 |
| 588 | 485.9 | 6.5 | 537 | 1 | AB051008 | ACCESSION:AB051008 |
| 589 | 485.9 | 6.5 | 537 | 1 | AB114927 | ACCESSION:AB114927 |
| 590 | 485.9 | 6.5 | 537 | 1 | AB114932 | ACCESSION:AB114932 |
| 591 | 485.9 | 6.5 | 537 | 1 | AB114938 | ACCESSION:AB114938 |
| 592 | 485.9 | 6.5 | 537 | 1 | AB114941 | ACCESSION:AB114941 |
| 593 | 485.9 | 6.5 | 537 | 1 | AB114974 | ACCESSION:AB114974 |
| 594 | 485.9 | 6.5 | 537 | 1 | AB114983 | ACCESSION:AB114983 |
| 595 | 485.9 | 6.5 | 537 | 1 | AB114994 | ACCESSION:AB114994 |
| 596 | 485.9 | 6.5 | 537 | 1 | AB115098 | ACCESSION:AB115098 |
| 597 | 485.9 | 6.5 | 537 | 1 | AB115101 | ACCESSION:AB115101 |
| 598 | 485.9 | 6.5 | 537 | 1 | AB115106 | ACCESSION:AB115106 |
| 599 | 485.9 | 6.5 | 537 | 1 | AB115109 | ACCESSION:AB115109 |
| 600 | 485.9 | 6.5 | 537 | 1 | AB115127 | ACCESSION:AB115127 |
| 601 | 485.9 | 6.5 | 537 | 1 | AB115131 | ACCESSION:AB115131 |
| 602 | 485.9 | 6.5 | 537 | 1 | AB115149 | ACCESSION:AB115149 |
| 603 | 485.9 | 6.5 | 537 | 1 | AB115155 | ACCESSION:AB115155 |
| 604 | 485.9 | 6.5 | 537 | 1 | AB115163 | ACCESSION:AB115163 |
| 605 | 485.7 | 6.5 | 539 | 1 | AB115095 | ACCESSION:AB115095 |
| 606 | 485.7 | 6.5 | 539 | 1 | AB115148 | ACCESSION:AB115148 |
| 607 | 485.6 | 6.5 | 540 | 1 | AB051001 | ACCESSION:AB051001 |
| 608 | 485.6 | 6.5 | 540 | 1 | AB114984 | ACCESSION:AB114984 |
| 609 | 485.6 | 6.5 | 540 | 1 | AB114992 | ACCESSION:AB114992 |
| 610 | 485.5 | 6.5 | 537 | 1 | AB114933 | ACCESSION:AB114933 |
| 611 | 485.5 | 6.5 | 537 | 1 | AB115147 | ACCESSION:AB115147 |
| 612 | 485.3 | 6.5 | 539 | 1 | AB114980 | ACCESSION:AB114980 |
| 613 | 485.3 | 6.5 | 539 | 1 | AB115158 | ACCESSION:AB115158 |
| 614 | 484.4 | 6.5 | 536 | 1 | AB115135 | ACCESSION:AB115135 |
| 615 | 484.3 | 6.5 | 537 | 1 | AB055413 | ACCESSION:AB055413 |
| 616 | 484.3 | 6.5 | 537 | 1 | AB055415 | ACCESSION:AB055415 |
| 617 | 484.3 | 6.5 | 537 | 1 | AB114950 | ACCESSION:AB114950 |
| 618 | 484.3 | 6.5 | 537 | 1 | AB115004 | ACCESSION:AB115004 |
| 619 | 484.3 | 6.5 | 537 | 1 | AB115105 | ACCESSION:AB115105 |
| 620 | 484.3 | 6.5 | 537 | 1 | AB115161 | ACCESSION:AB115161 |
| 621 | 484.3 | 6.5 | 537 | 1 | AB115165 | ACCESSION:AB115165 |
| 622 | 484.3 | 6.5 | 537 | 1 | AB115175 | ACCESSION:AB115175 |
| 623 | 484.3 | 6.5 | 537 | 1 | AB115176 | ACCESSION:AB115176 |
| 624 | 483.9 | 6.5 | 537 | 1 | AB115145 | ACCESSION:AB115145 |
| 625 | 483.8 | 6.5 | 538 | 1 | AB115010 | ACCESSION:AB115010 |
| 626 | 483.7 | 6.5 | 539 | 1 | AB054086 | ACCESSION:AB054086 |
| 627 | 483.7 | 6.5 | 539 | 1 | AB115005 | ACCESSION:AB115005 |
| 628 | 483.3 | 6.5 | 652 | 1 | AB115293 | ACCESSION:AB115293 |
| 629 | 483 | 6.4 | 538 | 1 | AB096034 | ACCESSION:AB096034 |
| 630 | 482.8 | 6.4 | 540 | 1 | AB051000 | ACCESSION:AB051000 |
| 631 | 482.7 | 6.4 | 537 | 1 | AB115002 | ACCESSION:AB115002 |
| 632 | 482.7 | 6.4 | 537 | 1 | AB115100 | ACCESSION:AB115100 |
| 633 | 482.7 | 6.4 | 537 | 1 | AB115115 | ACCESSION:AB115115 |
| 634 | 482.7 | 6.4 | 537 | 1 | AB115167 | ACCESSION:AB115167 |
| 635 | 482.7 | 6.4 | 545 | 1 | AB114964 | ACCESSION:AB114964 |
| 636 | 482.4 | 6.4 | 540 | 1 | AB050999 | ACCESSION:AB050999 |
| 637 | 482.4 | 6.4 | 540 | 1 | AB096038 | ACCESSION:AB096038 |
| 638 | 482.4 | 6.4 | 540 | 1 | AB115162 | ACCESSION:AB115162 |
| c 639 | 481.1 | 6.4 | 549 | 1 | CQ101066 | ACCESSION:CQ101066 |
| c 640 | 481.1 | 6.4 | 549 | 1 | CQ140060 | ACCESSION:CQ140060 |
| c 641 | 481.1 | 6.4 | 549 | 1 | CQ176211 | ACCESSION:CQ176211 |
| c 642 | 481.1 | 6.4 | 549 | 1 | CQ223406 | ACCESSION:CQ223406 |
| c 643 | 481.1 | 6.4 | 549 | 1 | CQ261379 | ACCESSION:CQ261379 |
| c 644 | 481.1 | 6.4 | 549 | 1 | CQ298746 | ACCESSION:CQ298746 |
| c 645 | 481.1 | 6.4 | 549 | 1 | CQ335530 | ACCESSION:CQ335530 |
| 646 | 480.7 | 6.4 | 537 | 1 | AB115096 | ACCESSION:AB115096 |
| 647 | 480.7 | 6.4 | 537 | 1 | AB115146 | ACCESSION:AB115146 |
| 648 | 480.7 | 6.4 | 537 | 1 | AB115157 | ACCESSION:AB115157 |
| 649 | 480.5 | 6.4 | 539 | 1 | AB114986 | ACCESSION:AB114986 |
| 650 | 480.5 | 6.4 | 539 | 1 | AB114997 | ACCESSION:AB114997 |
| 651 | 479.2 | 6.4 | 536 | 1 | AB115094 | ACCESSION:AB115094 |
| 652 | 478.9 | 6.4 | 539 | 1 | AB114988 | ACCESSION:AB114988 |
| 653 | 478.9 | 6.4 | 539 | 1 | AB114995 | ACCESSION:AB114995 |
| 654 | 478.3 | 6.4 | 541 | 1 | AB096028 | ACCESSION:AB096028 |
| 655 | 477.9 | 6.4 | 537 | 1 | AB115159 | ACCESSION:AB115159 |
| 656 | 477.8 | 6.4 | 660 | 1 | CQ746475 | ACCESSION:CQ746475 |
| 657 | 477.7 | 6.4 | 539 | 1 | AB115140 | ACCESSION:AB115140 |
| 658 | 477.6 | 6.4 | 536 | 1 | AB115108 | ACCESSION:AB115108 |
| 659 | 477.6 | 6.4 | 547 | 1 | AB096027 | ACCESSION:AB096027 |
| 660 | 477 | 6.4 | 538 | 1 | AB126140 | ACCESSION:AB126140 |
| 661 | 476.4 | 6.4 | 539 | 1 | AB115137 | ACCESSION:AB115137 |
| 662 | 476.3 | 6.4 | 537 | 1 | AB115174 | ACCESSION:AB115174 |
| 663 | 476 | 6.4 | 541 | 1 | AB097588 | ACCESSION:AB097588 |
| 664 | 475.9 | 6.4 | 537 | 1 | AB114998 | ACCESSION:AB114998 |
| 665 | 475.9 | 6.4 | 537 | 1 | AB115164 | ACCESSION:AB115164 |
| 666 | 475.6 | 6.4 | 536 | 1 | AB126158 | ACCESSION:AB126158 |
| 667 | 474.4 | 6.3 | 536 | 1 | AB115129 | ACCESSION:AB115129 |
| 668 | 474.3 | 6.3 | 537 | 1 | AB115107 | ACCESSION:AB115107 |
| 669 | 473.5 | 6.3 | 661 | 1 | AB115250 | ACCESSION:AB115250 |
| 670 | 473.4 | 6.3 | 538 | 1 | AB126152 | ACCESSION:AB126152 |
| 671 | 473.1 | 6.3 | 537 | 1 | AB114929 | ACCESSION:AB114929 |
| 672 | 473.1 | 6.3 | 537 | 1 | AB114993 | ACCESSION:AB114993 |
| 673 | 472.8 | 6.3 | 536 | 1 | AB115171 | ACCESSION:AB115171 |
| 674 | 472.5 | 6.3 | 539 | 1 | AB079505 | ACCESSION:AB079505 |
| 675 | 472.5 | 6.3 | 656 | 1 | AB115252 | ACCESSION:AB115252 |
| 676 | 472.4 | 6.3 | 536 | 1 | AB126147 | ACCESSION:AB126147 |
| 677 | 471.4 | 6.3 | 534 | 1 | AB115125 | ACCESSION:AB115125 |
| 678 | 471.2 | 6.3 | 536 | 1 | AB051002 | ACCESSION:AB051002 |
| 679 | 471.1 | 6.3 | 537 | 1 | AB114985 | ACCESSION:AB114985 |
| 680 | 470.4 | 6.3 | 543 | 1 | AB079508 | ACCESSION:AB079508 |
| 681 | 470.4 | 6.3 | 544 | 1 | AB115152 | ACCESSION:AB115152 |
| 682 | 469.7 | 6.3 | 535 | 1 | AB114971 | ACCESSION:AB114971 |
| 683 | 469.6 | 6.3 | 536 | 1 | AB115166 | ACCESSION:AB115166 |
| 684 | 469.5 | 6.3 | 537 | 1 | AB050996 | ACCESSION:AB050996 |
| 685 | 469.3 | 6.3 | 539 | 1 | AB126151 | ACCESSION:AB126151 |
| 686 | 469.2 | 6.3 | 536 | 1 | AB126159 | ACCESSION:AB126159 |
| 687 | 469.2 | 6.3 | 536 | 1 | AB126165 | ACCESSION:AB126165 |
| 688 | 468.6 | 6.3 | 652 | 1 | AB115251 | ACCESSION:AB115251 |
| 689 | 468.4 | 6.3 | 540 | 1 | AB126153 | ACCESSION:AB126153 |
| 690 | 468.1 | 6.2 | 539 | 1 | AB097585 | ACCESSION:AB097585 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 691 | 468 | 6.2 | 536 | 1 | AB115177 | ACCESSION:AB115177 |
| 692 | 467.6 | 6.2 | 535 | 1 | AB115178 | ACCESSION:AB115178 |
| 693 | 467.6 | 6.2 | 536 | 1 | AB126161 | ACCESSION:AB126161 |
| 694 | 467.4 | 6.2 | 644 | 1 | AB115265 | ACCESSION:AB115265 |
| 695 | 467.1 | 6.2 | 539 | 1 | AB097584 | ACCESSION:AB097584 |
| 696 | 465.5 | 6.2 | 533 | 1 | AB115114 | ACCESSION:AB115114 |
| 697 | 465.2 | 6.2 | 543 | 1 | AB096029 | ACCESSION:AB096029 |
| 698 | 464.4 | 6.2 | 536 | 1 | AB055414 | ACCESSION:AB055414 |
| 699 | 463.9 | 6.2 | 539 | 1 | AB071980 | ACCESSION:AB071980 |
| 700 | 463.9 | 6.2 | 539 | 1 | AB097577 | ACCESSION:AB097577 |
| 701 | 463.3 | 6.2 | 535 | 1 | AB114958 | ACCESSION:AB114958 |
| 702 | 461.7 | 6.2 | 535 | 1 | AB114940 | ACCESSION:AB114940 |
| 703 | 461.6 | 6.2 | 532 | 1 | AB050997 | ACCESSION:AB050997 |
| 704 | 461.2 | 6.2 | 540 | 1 | AB096033 | ACCESSION:AB096033 |
| 705 | 460.7 | 6.2 | 568 | 1 | CQ104355 | ACCESSION:CQ104355 |
| 706 | 459.6 | 6.1 | 536 | 1 | AB079506 | ACCESSION:AB079506 |
| 707 | 459.6 | 6.1 | 536 | 1 | AB096036 | ACCESSION:AB096036 |
| 708 | 458.5 | 6.1 | 539 | 1 | AB096030 | ACCESSION:AB096030 |
| 709 | 458.5 | 6.1 | 543 | 1 | AB071979 | ACCESSION:AB071979 |
| 710 | 458.5 | 6.1 | 543 | 1 | AB097576 | ACCESSION:AB097576 |
| 711 | 456.1 | 6.1 | 535 | 1 | AB071981 | ACCESSION:AB071981 |
| 712 | 456.1 | 6.1 | 535 | 1 | AB097578 | ACCESSION:AB097578 |
| 713 | 456.1 | 6.1 | 535 | 1 | AB126133 | ACCESSION:AB126133 |
| 714 | 456.1 | 6.1 | 535 | 1 | AB126136 | ACCESSION:AB126136 |
| 715 | 456.1 | 6.1 | 535 | 1 | AB126141 | ACCESSION:AB126141 |
| 716 | 456.1 | 6.1 | 535 | 1 | AB126145 | ACCESSION:AB126145 |
| 717 | 456.1 | 6.1 | 535 | 1 | AB126148 | ACCESSION:AB126148 |
| 718 | 455.7 | 6.1 | 539 | 1 | AB071982 | ACCESSION:AB071982 |
| 719 | 455.7 | 6.1 | 539 | 1 | AB097579 | ACCESSION:AB097579 |
| 720 | 455.1 | 6.1 | 535 | 1 | AB126155 | ACCESSION:AB126155 |
| 721 | 455.1 | 6.1 | 535 | 1 | AB126163 | ACCESSION:AB126163 |
| 722 | 454.8 | 6.1 | 536 | 1 | AB114936 | ACCESSION:AB114936 |
| 723 | 454.7 | 6.1 | 533 | 1 | AB054088 | ACCESSION:AB054088 |
| 724 | 454.5 | 6.1 | 535 | 1 | AB097582 | ACCESSION:AB097582 |
| 725 | 454.5 | 6.1 | 535 | 1 | AB126139 | ACCESSION:AB126139 |
| 726 | 454.5 | 6.1 | 535 | 1 | AB126154 | ACCESSION:AB126154 |
| 727 | 454.5 | 6.1 | 535 | 1 | AB126164 | ACCESSION:AB126164 |
| 728 | 454.1 | 6.1 | 539 | 1 | AB097580 | ACCESSION:AB097580 |
| c 729 | 454 | 6.1 | 548 | 1 | CQ101825 | ACCESSION:CQ101825 |
| c 730 | 454 | 6.1 | 548 | 1 | CQ140811 | ACCESSION:CQ140811 |
| c 731 | 454 | 6.1 | 548 | 1 | CQ224106 | ACCESSION:CQ224106 |
| c 732 | 454 | 6.1 | 548 | 1 | CQ262103 | ACCESSION:CQ262103 |
| c 733 | 454 | 6.1 | 548 | 1 | CQ299248 | ACCESSION:CQ299248 |
| c 734 | 454 | 6.1 | 548 | 1 | CQ336270 | ACCESSION:CQ336270 |
| 735 | 453.5 | 6.1 | 535 | 1 | AB126157 | ACCESSION:AB126157 |
| 736 | 452.9 | 6.0 | 535 | 1 | AB126150 | ACCESSION:AB126150 |
| 737 | 452.9 | 6.0 | 535 | 1 | AB126162 | ACCESSION:AB126162 |
| 738 | 452.3 | 6.0 | 526 | 1 | AB120060 | ACCESSION:AB120060 |
| 739 | 451.6 | 6.0 | 524 | 1 | AB097581 | ACCESSION:AB097581 |
| 740 | 451.6 | 6.0 | 524 | 1 | AB097586 | ACCESSION:AB097586 |
| 741 | 451.3 | 6.0 | 535 | 1 | AB126138 | ACCESSION:AB126138 |
| 742 | 451.2 | 6.0 | 544 | 1 | AB096032 | ACCESSION:AB096032 |
| 743 | 450.7 | 6.0 | 535 | 1 | AB126144 | ACCESSION:AB126144 |
| 744 | 450.7 | 6.0 | 537 | 1 | AB126134 | ACCESSION:AB126134 |
| 745 | 450.7 | 6.0 | 543 | 1 | AB096021 | ACCESSION:AB096021 |
| 746 | 449.7 | 6.0 | 535 | 1 | AB126146 | ACCESSION:AB126146 |
| 747 | 449.4 | 6.0 | 540 | 1 | AB079507 | ACCESSION:AB079507 |
| 748 | 448.7 | 6.0 | 535 | 1 | AB126143 | ACCESSION:AB126143 |
| 749 | 447.9 | 6.0 | 537 | 1 | AB126156 | ACCESSION:AB126156 |
| 750 | 445.5 | 5.9 | 525 | 1 | AB120059 | ACCESSION:AB120059 |
| c 751 | 445.1 | 5.9 | 559 | 1 | CQ071133 | ACCESSION:CQ071133 |
| c 752 | 445.1 | 5.9 | 559 | 1 | CQ100883 | ACCESSION:CQ100883 |
| c 753 | 445.1 | 5.9 | 559 | 1 | CQ139878 | ACCESSION:CQ139878 |
| c 754 | 445.1 | 5.9 | 559 | 1 | CQ176105 | ACCESSION:CQ176105 |
| c 755 | 445.1 | 5.9 | 559 | 1 | CQ223232 | ACCESSION:CQ223232 |
| c 756 | 445.1 | 5.9 | 559 | 1 | CQ261200 | ACCESSION:CQ261200 |
| c 757 | 445.1 | 5.9 | 559 | 1 | CQ298620 | ACCESSION:CQ298620 |
| c 758 | 445.1 | 5.9 | 559 | 1 | CQ335348 | ACCESSION:CQ335348 |
| 759 | 444.6 | 5.9 | 534 | 1 | AB097583 | ACCESSION:AB097583 |
| 760 | 444.3 | 5.9 | 541 | 1 | AB051006 | ACCESSION:AB051006 |
| 761 | 442.4 | 5.9 | 534 | 1 | AB096035 | ACCESSION:AB096035 |
| 762 | 442.1 | 5.9 | 539 | 1 | AB126135 | ACCESSION:AB126135 |
| 763 | 441.4 | 5.9 | 540 | 1 | AB097587 | ACCESSION:AB097587 |
| 764 | 440.7 | 5.9 | 549 | 1 | AB114965 | ACCESSION:AB114965 |
| 765 | 438.8 | 5.9 | 544 | 1 | AB096031 | ACCESSION:AB096031 |
| 766 | 437.4 | 5.8 | 601 | 1 | BV198554 | ACCESSION:BV198554 |
| 767 | 436.8 | 5.8 | 443 | 1 | AX905531 | ACCESSION:AX905531 |
| 768 | 436.8 | 5.8 | 443 | 1 | BD041064 | ACCESSION:BD041064 |
| 769 | 436.8 | 5.8 | 443 | 1 | AR741325 | ACCESSION:AR741325 |
| 770 | 433.2 | 5.8 | 491 | 1 | CQ109559 | ACCESSION:CQ109559 |
| 771 | 425.8 | 5.7 | 538 | 1 | AB126132 | ACCESSION:AB126132 |
| 772 | 423.1 | 5.6 | 549 | 1 | AB114955 | ACCESSION:AB114955 |
| 773 | 420.9 | 5.6 | 558 | 1 | CQ104018 | ACCESSION:CQ104018 |
| 774 | 415.4 | 5.5 | 542 | 1 | AB096037 | ACCESSION:AB096037 |
| 775 | 411.9 | 5.5 | 459 | 1 | CQ748604 | ACCESSION:CQ748604 |
| 776 | 411.1 | 5.5 | 492 | 1 | CQ750414 | ACCESSION:CQ750414 |
| 777 | 409.6 | 5.5 | 488 | 1 | CQ050965 | ACCESSION:CQ050965 |
| 778 | 409.6 | 5.5 | 488 | 1 | CQ065989 | ACCESSION:CQ065989 |
| 779 | 409.6 | 5.5 | 488 | 1 | CQ093006 | ACCESSION:CQ093006 |
| 780 | 409.6 | 5.5 | 488 | 1 | CQ131815 | ACCESSION:CQ131815 |
| 781 | 409.6 | 5.5 | 488 | 1 | CQ170381 | ACCESSION:CQ170381 |
| 782 | 409.6 | 5.5 | 488 | 1 | CQ199555 | ACCESSION:CQ199555 |
| 783 | 409.6 | 5.5 | 488 | 1 | CQ215005 | ACCESSION:CQ215005 |
| 784 | 409.6 | 5.5 | 488 | 1 | CQ253598 | ACCESSION:CQ253598 |
| 785 | 409.6 | 5.5 | 488 | 1 | CQ290692 | ACCESSION:CQ290692 |
| 786 | 409.6 | 5.5 | 488 | 1 | CQ327720 | ACCESSION:CQ327720 |
| c 787 | 400.4 | 5.3 | 504 | 1 | CQ099994 | ACCESSION:CQ099994 |
| c 788 | 400.4 | 5.3 | 504 | 1 | CQ138977 | ACCESSION:CQ138977 |
| c 789 | 400.4 | 5.3 | 504 | 1 | CQ175631 | ACCESSION:CQ175631 |
| c 790 | 400.4 | 5.3 | 504 | 1 | CQ222372 | ACCESSION:CQ222372 |
| c 791 | 400.4 | 5.3 | 504 | 1 | CQ260325 | ACCESSION:CQ260325 |
| c 792 | 400.4 | 5.3 | 504 | 1 | CQ298000 | ACCESSION:CQ298000 |
| c 793 | 400.4 | 5.3 | 504 | 1 | CQ334456 | ACCESSION:CQ334456 |
| 794 | 396.9 | 5.3 | 480 | 1 | CQ744629 | ACCESSION:CQ744629 |
| c 795 | 385.7 | 5.1 | 534 | 1 | AF328493 | ACCESSION:AF328493 |
| 796 | 384.5 | 5.1 | 444 | 1 | CQ752562 | ACCESSION:CQ752562 |
| 797 | 382.5 | 5.1 | 438 | 1 | AR344428 | ACCESSION:AR344428 |
| 798 | 381.3 | 5.1 | 438 | 1 | AR344405 | ACCESSION:AR344405 |
| 799 | 381.3 | 5.1 | 438 | 1 | AR344406 | ACCESSION:AR344406 |
| 800 | 381.3 | 5.1 | 438 | 1 | AR344408 | ACCESSION:AR344408 |
| 801 | 380.9 | 5.1 | 429 | 1 | AB024472 | ACCESSION:AB024472 |
| 802 | 380.5 | 5.1 | 429 | 1 | AB024473 | ACCESSION:AB024473 |
| 803 | 379.7 | 5.1 | 438 | 1 | AR344407 | ACCESSION:AR344407 |
| 804 | 378.9 | 5.1 | 429 | 1 | AB024469 | ACCESSION:AB024469 |
| 805 | 378.6 | 5.1 | 433 | 1 | CQ096442 | ACCESSION:CQ096442 |
| 806 | 378.3 | 5.1 | 429 | 1 | CQ747668 | ACCESSION:CQ747668 |
| 807 | 377.7 | 5.0 | 422 | 1 | CQ068125 | ACCESSION:CQ068125 |
| 808 | 377.7 | 5.0 | 422 | 1 | CQ095200 | ACCESSION:CQ095200 |
| 809 | 377.7 | 5.0 | 422 | 1 | CQ133948 | ACCESSION:CQ133948 |
| 810 | 377.7 | 5.0 | 422 | 1 | CQ201644 | ACCESSION:CQ201644 |
| 811 | 377.7 | 5.0 | 422 | 1 | CQ217189 | ACCESSION:CQ217189 |
| 812 | 377.3 | 5.0 | 429 | 1 | AB024484 | ACCESSION:AB024484 |
| 813 | 376.1 | 5.0 | 429 | 1 | AB024467 | ACCESSION:AB024467 |
| 814 | 375.7 | 5.0 | 429 | 1 | AB024471 | ACCESSION:AB024471 |
| 815 | 375.7 | 5.0 | 429 | 1 | AB024474 | ACCESSION:AB024474 |
| 816 | 375.5 | 5.0 | 429 | 1 | CQ751113 | ACCESSION:CQ751113 |
| 817 | 375.3 | 5.0 | 429 | 1 | AB024488 | ACCESSION:AB024488 |
| 818 | 374.8 | 5.0 | 492 | 1 | CQ742800 | ACCESSION:CQ742800 |
| 819 | 374.7 | 5.0 | 429 | 1 | AB024459 | ACCESSION:AB024459 |
| 820 | 374.5 | 5.0 | 429 | 1 | AB024480 | ACCESSION:AB024480 |
| 821 | 374.1 | 5.0 | 429 | 1 | AB024457 | ACCESSION:AB024457 |
| 822 | 373.1 | 5.0 | 429 | 1 | AB024460 | ACCESSION:AB024460 |
| 823 | 372.5 | 5.0 | 429 | 1 | AB024470 | ACCESSION:AB024470 |
| 824 | 372.1 | 5.0 | 429 | 1 | AB024487 | ACCESSION:AB024487 |
| 825 | 371.5 | 5.0 | 429 | 1 | AB024458 | ACCESSION:AB024458 |
| 826 | 370.9 | 5.0 | 429 | 1 | AB024462 | ACCESSION:AB024462 |
| 827 | 370.9 | 5.0 | 429 | 1 | AB024479 | ACCESSION:AB024479 |
| 828 | 370.5 | 4.9 | 429 | 1 | AB024461 | ACCESSION:AB024461 |
| 829 | 368 | 4.9 | 428 | 1 | AB024481 | ACCESSION:AB024481 |
| 830 | 367.7 | 4.9 | 429 | 1 | AB021921 | ACCESSION:AB021921 |
| 831 | 366.7 | 4.9 | 429 | 1 | AB024463 | ACCESSION:AB024463 |
| 832 | 366.7 | 4.9 | 429 | 1 | AB024464 | ACCESSION:AB024464 |
| 833 | 363.8 | 4.9 | 428 | 1 | AB024476 | ACCESSION:AB024476 |
| 834 | 362.2 | 4.8 | 426 | 1 | AB024465 | ACCESSION:AB024465 |
| 835 | 361.5 | 4.8 | 431 | 1 | AB024466 | ACCESSION:AB024466 |
| 836 | 359.5 | 4.8 | 427 | 1 | AB024486 | ACCESSION:AB024486 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 837 | 357.7 | 4.8 | 429 | 1 | AB024475 | ACCESSION:AB024475 |
| | 838 | 357.5 | 4.8 | 373 | 1 | AB055868 | ACCESSION:AB055868 |
| | 839 | 357.5 | 4.8 | 373 | 1 | AB055876 | ACCESSION:AB055876 |
| | 840 | 357.5 | 4.8 | 373 | 1 | AB064529 | ACCESSION:AB064529 |
| | 841 | 357.5 | 4.8 | 373 | 1 | AB064533 | ACCESSION:AB064533 |
| | 842 | 357.5 | 4.8 | 373 | 1 | AB064538 | ACCESSION:AB064538 |
| | 843 | 357.5 | 4.8 | 373 | 1 | AB064539 | ACCESSION:AB064539 |
| | 844 | 357.5 | 4.8 | 373 | 1 | AB066648 | ACCESSION:AB066648 |
| | 845 | 357.5 | 4.8 | 373 | 1 | AB066659 | ACCESSION:AB066659 |
| | 846 | 357.5 | 4.8 | 373 | 1 | AB071972 | ACCESSION:AB071972 |
| | 847 | 356.5 | 4.8 | 373 | 1 | AB066205 | ACCESSION:AB066205 |
| | 848 | 355.9 | 4.8 | 373 | 1 | AB055871 | ACCESSION:AB055871 |
| | 849 | 355.9 | 4.8 | 373 | 1 | AB055872 | ACCESSION:AB055872 |
| | 850 | 355.9 | 4.8 | 373 | 1 | AB055874 | ACCESSION:AB055874 |
| | 851 | 355.9 | 4.8 | 373 | 1 | AB064530 | ACCESSION:AB064530 |
| | 852 | 355.9 | 4.8 | 373 | 1 | AB064531 | ACCESSION:AB064531 |
| | 853 | 355.9 | 4.8 | 373 | 1 | AB064535 | ACCESSION:AB064535 |
| | 854 | 355.9 | 4.8 | 373 | 1 | AB064537 | ACCESSION:AB064537 |
| | 855 | 355.9 | 4.8 | 373 | 1 | AB066204 | ACCESSION:AB066204 |
| | 856 | 355.9 | 4.8 | 373 | 1 | AB066655 | ACCESSION:AB066655 |
| | 857 | 355.9 | 4.8 | 373 | 1 | AB066656 | ACCESSION:AB066656 |
| | 858 | 355.9 | 4.8 | 373 | 1 | AB066657 | ACCESSION:AB066657 |
| | 859 | 354.3 | 4.7 | 373 | 1 | AB055873 | ACCESSION:AB055873 |
| | 860 | 354.3 | 4.7 | 373 | 1 | AB055877 | ACCESSION:AB055877 |
| | 861 | 354.3 | 4.7 | 373 | 1 | AB064536 | ACCESSION:AB064536 |
| | 862 | 354.3 | 4.7 | 373 | 1 | AB064541 | ACCESSION:AB064541 |
| | 863 | 354.3 | 4.7 | 373 | 1 | AB064542 | ACCESSION:AB064542 |
| | 864 | 354.3 | 4.7 | 373 | 1 | AB066653 | ACCESSION:AB066653 |
| | 865 | 354.3 | 4.7 | 373 | 1 | AB066654 | ACCESSION:AB066654 |
| | 866 | 353.3 | 4.7 | 373 | 1 | AB066203 | ACCESSION:AB066203 |
| | 867 | 352.9 | 4.7 | 384 | 1 | BD058454 | ACCESSION:BD058454 |
| | 868 | 352.7 | 4.7 | 373 | 1 | AB064540 | ACCESSION:AB064540 |
| | 869 | 352.7 | 4.7 | 373 | 1 | AB066201 | ACCESSION:AB066201 |
| | 870 | 351.5 | 4.7 | 427 | 1 | AB024468 | ACCESSION:AB024468 |
| | 871 | 351.1 | 4.7 | 373 | 1 | AB064534 | ACCESSION:AB064534 |
| | 872 | 351.1 | 4.7 | 373 | 1 | AB066202 | ACCESSION:AB066202 |
| | 873 | 350.9 | 4.7 | 427 | 1 | AB024483 | ACCESSION:AB024483 |
| | 874 | 349.5 | 4.7 | 373 | 1 | AB079272 | ACCESSION:AB079272 |
| | 875 | 347.9 | 4.6 | 373 | 1 | AB066649 | ACCESSION:AB066649 |
| | 876 | 345.4 | 4.6 | 374 | 1 | AB071975 | ACCESSION:AB071975 |
| | 877 | 345 | 4.6 | 429 | 1 | AR344412 | ACCESSION:AR344412 |
| | 878 | 344.6 | 4.6 | 429 | 1 | AR344409 | ACCESSION:AR344409 |
| | 879 | 343 | 4.6 | 428 | 1 | AB024482 | ACCESSION:AB024482 |
| | 880 | 341.4 | 4.6 | 409 | 1 | CQ741551 | ACCESSION:CQ741551 |
| | 881 | 341 | 4.6 | 429 | 1 | AR344410 | ACCESSION:AR344410 |
| | 882 | 340.6 | 4.5 | 374 | 1 | AB055878 | ACCESSION:AB055878 |
| | 883 | 339.6 | 4.5 | 372 | 1 | AB066658 | ACCESSION:AB066658 |
| | 884 | 339.5 | 4.5 | 373 | 1 | AB079275 | ACCESSION:AB079275 |
| c | 885 | 339 | 4.5 | 410 | 1 | AX407891 | ACCESSION:AX407891 |
| | 886 | 339 | 4.5 | 429 | 1 | AR344411 | ACCESSION:AR344411 |
| | 887 | 338.3 | 4.5 | 373 | 1 | AB064532 | ACCESSION:AB064532 |
| | 888 | 338.3 | 4.5 | 373 | 1 | AB071973 | ACCESSION:AB071973 |
| | 889 | 337.9 | 4.5 | 389 | 1 | AB079279 | ACCESSION:AB079279 |
| | 890 | 336.3 | 4.5 | 389 | 1 | AB079269 | ACCESSION:AB079269 |
| | 891 | 336 | 4.5 | 434 | 1 | AB021920 | ACCESSION:AB021920 |
| c | 892 | 335.8 | 4.5 | 405 | 1 | CQ052505 | ACCESSION:CQ052505 |
| c | 893 | 335.8 | 4.5 | 405 | 1 | CQ067587 | ACCESSION:CQ067587 |
| c | 894 | 335.8 | 4.5 | 405 | 1 | CQ094637 | ACCESSION:CQ094637 |
| c | 895 | 335.8 | 4.5 | 405 | 1 | CQ133399 | ACCESSION:CQ133399 |
| c | 896 | 335.8 | 4.5 | 405 | 1 | CQ171968 | ACCESSION:CQ171968 |
| c | 897 | 335.8 | 4.5 | 405 | 1 | CQ201114 | ACCESSION:CQ201114 |
| c | 898 | 335.8 | 4.5 | 405 | 1 | CQ216631 | ACCESSION:CQ216631 |
| c | 899 | 335.8 | 4.5 | 405 | 1 | CQ255195 | ACCESSION:CQ255195 |
| c | 900 | 335.8 | 4.5 | 405 | 1 | CQ292282 | ACCESSION:CQ292282 |
| c | 901 | 335.8 | 4.5 | 405 | 1 | CQ329310 | ACCESSION:CQ329310 |
| | 902 | 335.7 | 4.5 | 373 | 1 | AB079273 | ACCESSION:AB079273 |
| c | 903 | 335.1 | 4.5 | 401 | 1 | BV193474 | ACCESSION:BV193474 |
| | 904 | 334.1 | 4.5 | 426 | 1 | AB024477 | ACCESSION:AB024477 |
| | 905 | 333.5 | 4.5 | 415 | 1 | AB024478 | ACCESSION:AB024478 |
| | 906 | 326.4 | 4.4 | 434 | 1 | AB021919 | ACCESSION:AB021919 |
| c | 907 | 324.7 | 4.3 | 409 | 1 | CQ053304 | ACCESSION:CQ053304 |
| c | 908 | 324.7 | 4.3 | 409 | 1 | CQ068432 | ACCESSION:CQ068432 |
| c | 909 | 324.7 | 4.3 | 409 | 1 | CQ095502 | ACCESSION:CQ095502 |
| c | 910 | 324.7 | 4.3 | 409 | 1 | CQ134252 | ACCESSION:CQ134252 |
| c | 911 | 324.7 | 4.3 | 409 | 1 | CQ172757 | ACCESSION:CQ172757 |
| c | 912 | 324.7 | 4.3 | 409 | 1 | CQ201948 | ACCESSION:CQ201948 |
| c | 913 | 324.7 | 4.3 | 409 | 1 | CQ217505 | ACCESSION:CQ217505 |
| c | 914 | 324.7 | 4.3 | 409 | 1 | CQ256070 | ACCESSION:CQ256070 |
| c | 915 | 324.7 | 4.3 | 409 | 1 | CQ293157 | ACCESSION:CQ293157 |
| c | 916 | 324.7 | 4.3 | 409 | 1 | CQ330127 | ACCESSION:CQ330127 |
| c | 917 | 324.3 | 4.3 | 401 | 1 | BV193437 | ACCESSION:BV193437 |
| c | 918 | 324.1 | 4.3 | 369 | 1 | G34081 | ACCESSION:G34081 |
| | 919 | 321.5 | 4.3 | 401 | 1 | BV193111 | ACCESSION:BV193111 |
| | 920 | 320.9 | 4.3 | 389 | 1 | AB066652 | ACCESSION:AB066652 |
| | 921 | 320.3 | 4.3 | 389 | 1 | AB079282 | ACCESSION:AB079282 |
| | 922 | 318.3 | 4.2 | 389 | 1 | AB066650 | ACCESSION:AB066650 |
| | 923 | 317.7 | 4.2 | 389 | 1 | AB066651 | ACCESSION:AB066651 |
| | 924 | 316.5 | 4.2 | 372 | 1 | AB079283 | ACCESSION:AB079283 |
| | 925 | 316.4 | 4.2 | 376 | 1 | AB055870 | ACCESSION:AB055870 |
| | 926 | 315.5 | 4.2 | 389 | 1 | AB079281 | ACCESSION:AB079281 |
| c | 927 | 315.5 | 4.2 | 414 | 1 | CQ053375 | ACCESSION:CQ053375 |
| c | 928 | 315.5 | 4.2 | 414 | 1 | CQ068513 | ACCESSION:CQ068513 |
| c | 929 | 315.5 | 4.2 | 414 | 1 | CQ095583 | ACCESSION:CQ095583 |
| c | 930 | 315.5 | 4.2 | 414 | 1 | CQ134328 | ACCESSION:CQ134328 |
| c | 931 | 315.5 | 4.2 | 414 | 1 | CQ172831 | ACCESSION:CQ172831 |
| c | 932 | 315.5 | 4.2 | 414 | 1 | CQ202022 | ACCESSION:CQ202022 |
| c | 933 | 315.5 | 4.2 | 414 | 1 | CQ217589 | ACCESSION:CQ217589 |
| c | 934 | 315.5 | 4.2 | 414 | 1 | CQ256153 | ACCESSION:CQ256153 |
| c | 935 | 315.5 | 4.2 | 414 | 1 | CQ293239 | ACCESSION:CQ293239 |
| c | 936 | 315.5 | 4.2 | 414 | 1 | CQ330194 | ACCESSION:CQ330194 |
| | 937 | 312.8 | 4.2 | 357 | 1 | CQ750358 | ACCESSION:CQ750358 |
| | 938 | 309.4 | 4.1 | 390 | 1 | AB079276 | ACCESSION:AB079276 |
| | 939 | 309.4 | 4.1 | 392 | 1 | AB079280 | ACCESSION:AB079280 |
| | 940 | 308.2 | 4.1 | 389 | 1 | AB079278 | ACCESSION:AB079278 |
| c | 941 | 307.6 | 4.1 | 401 | 1 | BV190586 | ACCESSION:BV190586 |
| | 942 | 306.3 | 4.1 | 374 | 1 | BD003152 | ACCESSION:BD003152 |
| | 943 | 306.3 | 4.1 | 374 | 1 | I80071 | ACCESSION:I80071 |
| | 944 | 300.1 | 4.0 | 375 | 1 | AB071970 | ACCESSION:AB071970 |
| | 945 | 295.7 | 3.9 | 372 | 1 | AB079294 | ACCESSION:AB079294 |
| | 946 | 295.1 | 3.9 | 373 | 1 | AB079274 | ACCESSION:AB079274 |
| c | 947 | 293.8 | 3.9 | 401 | 1 | BV195837 | ACCESSION:BV195837 |
| | 948 | 293.5 | 3.9 | 373 | 1 | AB071974 | ACCESSION:AB071974 |
| c | 949 | 292.1 | 3.9 | 336 | 1 | CQ114106 | ACCESSION:CQ114106 |
| c | 950 | 292.1 | 3.9 | 336 | 1 | CQ152987 | ACCESSION:CQ152987 |
| c | 951 | 292.1 | 3.9 | 336 | 1 | CQ186062 | ACCESSION:CQ186062 |
| c | 952 | 292.1 | 3.9 | 336 | 1 | CQ236297 | ACCESSION:CQ236297 |
| c | 953 | 292.1 | 3.9 | 336 | 1 | CQ273879 | ACCESSION:CQ273879 |
| c | 954 | 292.1 | 3.9 | 336 | 1 | CQ311149 | ACCESSION:CQ311149 |
| c | 955 | 292.1 | 3.9 | 336 | 1 | CQ348183 | ACCESSION:CQ348183 |
| | 956 | 291.5 | 3.9 | 372 | 1 | AB079291 | ACCESSION:AB079291 |
| | 957 | 291.5 | 3.9 | 373 | 1 | AB079277 | ACCESSION:AB079277 |
| | 958 | 289.9 | 3.9 | 372 | 1 | AB079286 | ACCESSION:AB079286 |
| | 959 | 288.7 | 3.9 | 373 | 1 | AB079296 | ACCESSION:AB079296 |
| | 960 | 288.6 | 3.9 | 329 | 1 | AY208135 | ACCESSION:AY208135 |
| | 961 | 287 | 3.8 | 329 | 1 | AY208134 | ACCESSION:AY208134 |
| c | 962 | 285.7 | 3.8 | 387 | 1 | CQ081963 | ACCESSION:CQ081963 |
| c | 963 | 285.7 | 3.8 | 387 | 1 | CQ116633 | ACCESSION:CQ116633 |
| c | 964 | 285.7 | 3.8 | 387 | 1 | CQ238553 | ACCESSION:CQ238553 |
| c | 965 | 285.7 | 3.8 | 387 | 1 | CQ276189 | ACCESSION:CQ276189 |
| c | 966 | 285.7 | 3.8 | 387 | 1 | CQ313183 | ACCESSION:CQ313183 |
| c | 967 | 285.7 | 3.8 | 387 | 1 | CQ350706 | ACCESSION:CQ350706 |
| | 968 | 285.4 | 3.8 | 329 | 1 | AY208132 | ACCESSION:AY208132 |
| | 969 | 283.3 | 3.8 | 341 | 1 | BD268834 | ACCESSION:BD268834 |
| | 970 | 283.3 | 3.8 | 341 | 1 | AX035742 | ACCESSION:AX035742 |
| | 971 | 281.4 | 3.8 | 375 | 1 | AB079270 | ACCESSION:AB079270 |
| | 972 | 281 | 3.8 | 374 | 1 | AB055869 | ACCESSION:AB055869 |
| | 973 | 280.6 | 3.7 | 328 | 1 | AY208133 | ACCESSION:AY208133 |
| | 974 | 280.4 | 3.7 | 374 | 1 | AB055875 | ACCESSION:AB055875 |
| | 975 | 280.4 | 3.7 | 374 | 1 | AB055879 | ACCESSION:AB055879 |
| | 976 | 277.8 | 3.7 | 371 | 1 | AB079284 | ACCESSION:AB079284 |
| | 977 | 276.5 | 3.7 | 346 | 1 | AF531174 | ACCESSION:AF531174 |
| | 978 | 274.6 | 3.7 | 372 | 1 | AB079288 | ACCESSION:AB079288 |
| | 979 | 273.8 | 3.7 | 373 | 1 | AB071971 | ACCESSION:AB071971 |
| | 980 | 271.7 | 3.6 | 373 | 1 | AB079287 | ACCESSION:AB079287 |
| | 981 | 270.3 | 3.6 | 343 | 1 | BD268847 | ACCESSION:BD268847 |
| | 982 | 270.3 | 3.6 | 343 | 1 | AX035755 | ACCESSION:AX035755 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 983 | 269.9 | 3.6 | 343 | 1 | BD268845 | ACCESSION:BD268845 |
| 984 | 269.9 | 3.6 | 343 | 1 | AX035753 | ACCESSION:AX035753 |
| c 985 | 268.4 | 3.6 | 379 | 1 | CQ052449 | ACCESSION:CQ052449 |
| c 986 | 268.4 | 3.6 | 379 | 1 | CQ067518 | ACCESSION:CQ067518 |
| c 987 | 268.4 | 3.6 | 379 | 1 | CQ094566 | ACCESSION:CQ094566 |
| c 988 | 268.4 | 3.6 | 379 | 1 | CQ133333 | ACCESSION:CQ133333 |
| c 989 | 268.4 | 3.6 | 379 | 1 | CQ201049 | ACCESSION:CQ201049 |
| c 990 | 268.4 | 3.6 | 379 | 1 | CQ216554 | ACCESSION:CQ216554 |
| c 991 | 268.4 | 3.6 | 379 | 1 | CQ255127 | ACCESSION:CQ255127 |
| c 992 | 268.4 | 3.6 | 379 | 1 | CQ292212 | ACCESSION:CQ292212 |
| c 993 | 268.4 | 3.6 | 379 | 1 | CQ329245 | ACCESSION:CQ329245 |
| 994 | 265.1 | 3.5 | 343 | 1 | BD268841 | ACCESSION:BD268841 |
| 995 | 265.1 | 3.5 | 343 | 1 | AX035749 | ACCESSION:AX035749 |
| 996 | 263.4 | 3.5 | 328 | 1 | AY208136 | ACCESSION:AY208136 |
| 997 | 262.9 | 3.5 | 341 | 1 | BD268835 | ACCESSION:BD268835 |
| 998 | 262.9 | 3.5 | 341 | 1 | AX035743 | ACCESSION:AX035743 |
| 999 | 262.7 | 3.5 | 310 | 1 | CQ055957 | ACCESSION:CQ055957 |
| 1000 | 262.7 | 3.5 | 310 | 1 | CQ075199 | ACCESSION:CQ075199 |
| 1001 | 262.7 | 3.5 | 310 | 1 | CQ106152 | ACCESSION:CQ106152 |
| 1002 | 262.7 | 3.5 | 310 | 1 | CQ144842 | ACCESSION:CQ144842 |
| 1003 | 262.7 | 3.5 | 310 | 1 | CQ180280 | ACCESSION:CQ180280 |
| 1004 | 262.7 | 3.5 | 310 | 1 | CQ204678 | ACCESSION:CQ204678 |
| 1005 | 262.7 | 3.5 | 310 | 1 | CQ228028 | ACCESSION:CQ228028 |
| 1006 | 262.7 | 3.5 | 310 | 1 | CQ266186 | ACCESSION:CQ266186 |
| 1007 | 262.7 | 3.5 | 310 | 1 | CQ303217 | ACCESSION:CQ303217 |
| 1008 | 262.7 | 3.5 | 310 | 1 | CQ340460 | ACCESSION:CQ340460 |
| 1009 | 260.3 | 3.5 | 343 | 1 | BD268839 | ACCESSION:BD268839 |
| 1010 | 260.3 | 3.5 | 343 | 1 | BD268844 | ACCESSION:BD268844 |
| 1011 | 260.3 | 3.5 | 343 | 1 | AX035747 | ACCESSION:AX035747 |
| 1012 | 260.3 | 3.5 | 343 | 1 | AX035752 | ACCESSION:AX035752 |
| 1013 | 258.7 | 3.5 | 343 | 1 | BD268840 | ACCESSION:BD268840 |
| 1014 | 258.7 | 3.5 | 343 | 1 | AX035748 | ACCESSION:AX035748 |
| 1015 | 255.5 | 3.4 | 343 | 1 | BD268842 | ACCESSION:BD268842 |
| 1016 | 255.5 | 3.4 | 343 | 1 | BD268843 | ACCESSION:BD268843 |
| 1017 | 255.5 | 3.4 | 343 | 1 | AX035750 | ACCESSION:AX035750 |
| 1018 | 255.5 | 3.4 | 343 | 1 | AX035751 | ACCESSION:AX035751 |
| 1019 | 253.9 | 3.4 | 343 | 1 | BD268837 | ACCESSION:BD268837 |
| 1020 | 253.9 | 3.4 | 343 | 1 | AX035745 | ACCESSION:AX035745 |
| 1021 | 252.3 | 3.4 | 343 | 1 | BD268846 | ACCESSION:BD268846 |
| 1022 | 252.3 | 3.4 | 343 | 1 | AX035754 | ACCESSION:AX035754 |
| 1023 | 250.7 | 3.3 | 343 | 1 | BD268838 | ACCESSION:BD268838 |
| 1024 | 250.7 | 3.3 | 343 | 1 | AX035746 | ACCESSION:AX035746 |
| c1025 | 246 | 3.3 | 300 | 1 | CQ741068 | ACCESSION:CQ741068 |
| 1026 | 244.5 | 3.3 | 275 | 1 | CQ077341 | ACCESSION:CQ077341 |
| 1027 | 244.5 | 3.3 | 275 | 1 | CQ108356 | ACCESSION:CQ108356 |
| 1028 | 244.5 | 3.3 | 275 | 1 | CQ146988 | ACCESSION:CQ146988 |
| 1029 | 244.5 | 3.3 | 275 | 1 | CQ206780 | ACCESSION:CQ206780 |
| 1030 | 244.5 | 3.3 | 275 | 1 | CQ230216 | ACCESSION:CQ230216 |
| c1031 | 243.1 | 3.2 | 299 | 1 | AF315117 | ACCESSION:AF315117 |
| c1032 | 242.7 | 3.2 | 299 | 1 | AF315115 | ACCESSION:AF315115 |
| 1033 | 238 | 3.2 | 279 | 1 | BD003154 | ACCESSION:BD003154 |
| 1034 | 238 | 3.2 | 279 | 1 | I80073 | ACCESSION:I80073 |
| c1035 | 237.9 | 3.2 | 299 | 1 | AF315111 | ACCESSION:AF315111 |
| 1036 | 237.1 | 3.2 | 310 | 1 | BD136194 | ACCESSION:BD136194 |
| 1037 | 237.1 | 3.2 | 310 | 1 | CS248636 | ACCESSION:CS248636 |
| 1038 | 237.1 | 3.2 | 310 | 1 | AR344384 | ACCESSION:AR344384 |
| 1039 | 237.1 | 3.2 | 310 | 1 | AX001022 | ACCESSION:AX001022 |
| 1040 | 235.6 | 3.1 | 290 | 1 | CQ686782 | ACCESSION:CQ686782 |
| 1041 | 234.4 | 3.1 | 322 | 1 | BD268836 | ACCESSION:BD268836 |
| 1042 | 234.4 | 3.1 | 322 | 1 | AX035744 | ACCESSION:AX035744 |
| c1043 | 233.1 | 3.1 | 299 | 1 | AF315109 | ACCESSION:AF315109 |
| c1044 | 233.1 | 3.1 | 299 | 1 | AF315114 | ACCESSION:AF315114 |
| c1045 | 231.5 | 3.1 | 299 | 1 | AF315110 | ACCESSION:AF315110 |
| 1046 | 228.3 | 3.0 | 288 | 1 | AY098484 | ACCESSION:AY098484 |
| c1047 | 228.3 | 3.0 | 299 | 1 | AF315112 | ACCESSION:AF315112 |
| c1048 | 228.3 | 3.0 | 299 | 1 | AF315113 | ACCESSION:AF315113 |
| c1049 | 226.7 | 3.0 | 299 | 1 | AF315107 | ACCESSION:AF315107 |
| c1050 | 225.1 | 3.0 | 299 | 1 | AF315116 | ACCESSION:AF315116 |
| c1051 | 223.5 | 3.0 | 299 | 1 | AF315108 | ACCESSION:AF315108 |
| 1052 | 221.7 | 3.0 | 258 | 1 | CQ739149 | ACCESSION:CQ739149 |
| 1053 | 218.6 | 2.9 | 249 | 1 | CQ747660 | ACCESSION:CQ747660 |
| 1054 | 215.6 | 2.9 | 224 | 1 | AY102609 | ACCESSION:AY102609 |
| 1055 | 215.2 | 2.9 | 297 | 1 | A46013 | ACCESSION:A46013 |
| 1056 | 215.2 | 2.9 | 297 | 1 | A57049 | ACCESSION:A57049 |
| 1057 | 215.2 | 2.9 | 297 | 1 | A60114 | ACCESSION:A60114 |
| 1058 | 215.2 | 2.9 | 297 | 1 | A79464 | ACCESSION:A79464 |
| 1059 | 215.2 | 2.9 | 297 | 1 | A80198 | ACCESSION:A80198 |
| 1060 | 215.2 | 2.9 | 297 | 1 | A80307 | ACCESSION:A80307 |
| 1061 | 215.2 | 2.9 | 297 | 1 | AR035350 | ACCESSION:AR035350 |
| 1062 | 215.2 | 2.9 | 297 | 1 | AR036191 | ACCESSION:AR036191 |
| 1063 | 215.2 | 2.9 | 297 | 1 | AR036711 | ACCESSION:AR036711 |
| 1064 | 215.2 | 2.9 | 297 | 1 | AR076973 | ACCESSION:AR076973 |
| 1065 | 215.2 | 2.9 | 297 | 1 | AR094694 | ACCESSION:AR094694 |
| 1066 | 215.2 | 2.9 | 297 | 1 | AR183929 | ACCESSION:AR183929 |
| 1067 | 215.2 | 2.9 | 297 | 1 | AR343141 | ACCESSION:AR343141 |
| 1068 | 215.2 | 2.9 | 297 | 1 | AR344298 | ACCESSION:AR344298 |
| 1069 | 207.4 | 2.8 | 211 | 1 | AX806706 | ACCESSION:AX806706 |
| 1070 | 205 | 2.7 | 278 | 1 | CQ055177 | ACCESSION:CQ055177 |
| 1071 | 205 | 2.7 | 278 | 1 | CQ074436 | ACCESSION:CQ074436 |
| 1072 | 205 | 2.7 | 278 | 1 | CQ105334 | ACCESSION:CQ105334 |
| 1073 | 205 | 2.7 | 278 | 1 | CQ144028 | ACCESSION:CQ144028 |
| 1074 | 205 | 2.7 | 278 | 1 | CQ179508 | ACCESSION:CQ179508 |
| 1075 | 205 | 2.7 | 278 | 1 | CQ203867 | ACCESSION:CQ203867 |
| 1076 | 205 | 2.7 | 278 | 1 | CQ227235 | ACCESSION:CQ227235 |
| 1077 | 205 | 2.7 | 278 | 1 | CQ265384 | ACCESSION:CQ265384 |
| 1078 | 205 | 2.7 | 278 | 1 | CQ302470 | ACCESSION:CQ302470 |
| 1079 | 205 | 2.7 | 278 | 1 | CQ339648 | ACCESSION:CQ339648 |
| 1080 | 199 | 2.7 | 201 | 1 | BV204004 | ACCESSION:BV204004 |
| 1081 | 199 | 2.7 | 201 | 1 | BV204007 | ACCESSION:BV204007 |
| 1082 | 199 | 2.7 | 201 | 1 | BV205765 | ACCESSION:BV205765 |
| 1083 | 199 | 2.7 | 201 | 1 | BV205767 | ACCESSION:BV205767 |
| 1084 | 199 | 2.7 | 201 | 1 | BV205773 | ACCESSION:BV205773 |
| 1085 | 198.6 | 2.7 | 201 | 1 | BV204005 | ACCESSION:BV204005 |
| 1086 | 193.9 | 2.6 | 240 | 1 | G01776 | ACCESSION:G01776 |
| 1087 | 189.6 | 2.5 | 201 | 1 | BV205774 | ACCESSION:BV205774 |
| 1088 | 178.6 | 2.4 | 201 | 1 | BV205775 | ACCESSION:BV205775 |
| c1089 | 177.2 | 2.4 | 200 | 1 | CQ058368 | ACCESSION:CQ058368 |
| c1090 | 177.2 | 2.4 | 200 | 1 | CQ077714 | ACCESSION:CQ077714 |
| c1091 | 177.2 | 2.4 | 200 | 1 | CQ108723 | ACCESSION:CQ108723 |
| c1092 | 177.2 | 2.4 | 200 | 1 | CQ147355 | ACCESSION:CQ147355 |
| c1093 | 177.2 | 2.4 | 200 | 1 | CQ182729 | ACCESSION:CQ182729 |
| c1094 | 177.2 | 2.4 | 200 | 1 | CQ207144 | ACCESSION:CQ207144 |
| c1095 | 177.2 | 2.4 | 200 | 1 | CQ230600 | ACCESSION:CQ230600 |
| c1096 | 177.2 | 2.4 | 200 | 1 | CQ268738 | ACCESSION:CQ268738 |
| c1097 | 177.2 | 2.4 | 200 | 1 | CQ305761 | ACCESSION:CQ305761 |
| c1098 | 177.2 | 2.4 | 200 | 1 | CQ342933 | ACCESSION:CQ342933 |
| 1099 | 172.2 | 2.3 | 217 | 1 | G01772 | ACCESSION:G01772 |
| 1100 | 165.8 | 2.2 | 201 | 1 | BV173708 | ACCESSION:BV173708 |
| 1101 | 165.6 | 2.2 | 198 | 1 | G30805 | ACCESSION:G30805 |
| c1102 | 165.6 | 2.2 | 222 | 1 | CQ746872 | ACCESSION:CQ746872 |
| 1103 | 162.8 | 2.2 | 201 | 1 | BV172221 | ACCESSION:BV172221 |
| c1104 | 158.2 | 2.1 | 201 | 1 | BV174185 | ACCESSION:BV174185 |
| c1105 | 157.6 | 2.1 | 222 | 1 | CQ745979 | ACCESSION:CQ745979 |
| 1106 | 156.6 | 2.1 | 202 | 1 | CQ116725 | ACCESSION:CQ116725 |
| 1107 | 156.4 | 2.1 | 201 | 1 | BV178412 | ACCESSION:BV178412 |
| 1108 | 156.2 | 2.1 | 207 | 1 | CQ106174 | ACCESSION:CQ106174 |
| 1109 | 156.2 | 2.1 | 207 | 1 | CQ228049 | ACCESSION:CQ228049 |
| 1110 | 148.9 | 2.0 | 190 | 1 | AY098483 | ACCESSION:AY098483 |
| 1111 | 139.6 | 1.9 | 165 | 1 | AY098485 | ACCESSION:AY098485 |
| 1112 | 139.2 | 1.9 | 150 | 1 | A60152 | ACCESSION:A60152 |
| 1113 | 139.2 | 1.9 | 150 | 1 | A79502 | ACCESSION:A79502 |
| 1114 | 139.2 | 1.9 | 150 | 1 | AR344331 | ACCESSION:AR344331 |
| c1115 | 137.7 | 1.8 | 174 | 1 | CQ743973 | ACCESSION:CQ743973 |
| c1116 | 136.9 | 1.8 | 180 | 1 | CQ739060 | ACCESSION:CQ739060 |
| c1117 | 132.8 | 1.8 | 186 | 1 | CQ752437 | ACCESSION:CQ752437 |
| c1118 | 127.3 | 1.7 | 180 | 1 | CQ742464 | ACCESSION:CQ742464 |
| c1119 | 126.7 | 1.7 | 172 | 1 | CQ740255 | ACCESSION:CQ740255 |
| 1120 | 122.6 | 1.6 | 152 | 1 | CQ058491 | ACCESSION:CQ058491 |
| 1121 | 122.6 | 1.6 | 152 | 1 | CQ077834 | ACCESSION:CQ077834 |
| 1122 | 122.6 | 1.6 | 152 | 1 | CQ108842 | ACCESSION:CQ108842 |
| 1123 | 122.6 | 1.6 | 152 | 1 | CQ147474 | ACCESSION:CQ147474 |
| 1124 | 122.6 | 1.6 | 152 | 1 | CQ207262 | ACCESSION:CQ207262 |
| 1125 | 122.6 | 1.6 | 152 | 1 | CQ230720 | ACCESSION:CQ230720 |
| 1126 | 122.6 | 1.6 | 152 | 1 | CQ268859 | ACCESSION:CQ268859 |
| 1127 | 122.6 | 1.6 | 152 | 1 | CQ305886 | ACCESSION:CQ305886 |
| 1128 | 121.8 | 1.6 | 133 | 1 | CQ117238 | ACCESSION:CQ117238 |

| | | | | | | |
|---|---|---|---|---|---|---|
| c1129 | 121.4 | 1.6 | 168 | 1 | CQ745278 | ACCESSION:CQ745278 |
| 1130 | 121 | 1.6 | 155 | 1 | CQ075479 | ACCESSION:CQ075479 |
| 1131 | 121 | 1.6 | 155 | 1 | CQ106452 | ACCESSION:CQ106452 |
| 1132 | 121 | 1.6 | 155 | 1 | CQ145113 | ACCESSION:CQ145113 |
| 1133 | 121 | 1.6 | 155 | 1 | CQ180556 | ACCESSION:CQ180556 |
| 1134 | 121 | 1.6 | 155 | 1 | CQ228295 | ACCESSION:CQ228295 |
| 1135 | 121 | 1.6 | 155 | 1 | CQ303441 | ACCESSION:CQ303441 |
| 1136 | 121 | 1.6 | 155 | 1 | CQ340739 | ACCESSION:CQ340739 |
| c1137 | 120.6 | 1.6 | 171 | 1 | CQ748116 | ACCESSION:CQ748116 |
| 1138 | 119.3 | 1.6 | 149 | 1 | CQ742851 | ACCESSION:CQ742851 |
| 1139 | 119.3 | 1.6 | 149 | 1 | CQ748132 | ACCESSION:CQ748132 |
| 1140 | 110.6 | 1.5 | 118 | 1 | AF009667 | ACCESSION:AF009667 |
| c1141 | 109.4 | 1.5 | 139 | 1 | HUMSWX525 | ACCESSION:L24603 |
| c1142 | 101.8 | 1.4 | 135 | 1 | CQ082276 | ACCESSION:CQ082276 |
| c1143 | 101.8 | 1.4 | 135 | 1 | CQ117002 | ACCESSION:CQ117002 |
| c1144 | 101.8 | 1.4 | 135 | 1 | CQ155677 | ACCESSION:CQ155677 |
| c1145 | 101.8 | 1.4 | 135 | 1 | CQ188073 | ACCESSION:CQ188073 |
| c1146 | 101.8 | 1.4 | 135 | 1 | CQ238853 | ACCESSION:CQ238853 |
| c1147 | 101.8 | 1.4 | 135 | 1 | CQ276533 | ACCESSION:CQ276533 |
| c1148 | 101.8 | 1.4 | 135 | 1 | CQ313549 | ACCESSION:CQ313549 |
| c1149 | 101.8 | 1.4 | 135 | 1 | CQ351031 | ACCESSION:CQ351031 |
| c1150 | 100.4 | 1.3 | 113 | 1 | CQ751084 | ACCESSION:CQ751084 |
| c1151 | 93 | 1.2 | 110 | 1 | CQ113073 | ACCESSION:CQ113073 |
| c1152 | 93 | 1.2 | 110 | 1 | CQ151943 | ACCESSION:CQ151943 |
| c1153 | 93 | 1.2 | 110 | 1 | CQ185508 | ACCESSION:CQ185508 |
| c1154 | 93 | 1.2 | 110 | 1 | CQ235300 | ACCESSION:CQ235300 |
| c1155 | 93 | 1.2 | 110 | 1 | CQ272864 | ACCESSION:CQ272864 |
| c1156 | 93 | 1.2 | 110 | 1 | CQ310432 | ACCESSION:CQ310432 |
| c1157 | 93 | 1.2 | 110 | 1 | CQ347148 | ACCESSION:CQ347148 |
| 1158 | 92 | 1.2 | 126 | 1 | A57086 | ACCESSION:A57086 |
| 1159 | 92 | 1.2 | 126 | 1 | A80344 | ACCESSION:A80344 |
| c1160 | 89.2 | 1.2 | 103 | 1 | CQ057462 | ACCESSION:CQ057462 |
| c1161 | 89.2 | 1.2 | 103 | 1 | CQ076748 | ACCESSION:CQ076748 |
| c1162 | 89.2 | 1.2 | 103 | 1 | CQ107735 | ACCESSION:CQ107735 |
| c1163 | 89.2 | 1.2 | 103 | 1 | CQ146384 | ACCESSION:CQ146384 |
| c1164 | 89.2 | 1.2 | 103 | 1 | CQ206196 | ACCESSION:CQ206196 |
| c1165 | 89.2 | 1.2 | 103 | 1 | CQ229597 | ACCESSION:CQ229597 |
| c1166 | 89.2 | 1.2 | 103 | 1 | CQ267736 | ACCESSION:CQ267736 |
| c1167 | 89.2 | 1.2 | 103 | 1 | CQ304760 | ACCESSION:CQ304760 |
| c1168 | 89.2 | 1.2 | 103 | 1 | CQ342005 | ACCESSION:CQ342005 |
| 1169 | 85.2 | 1.1 | 107 | 1 | CQ055176 | ACCESSION:CQ055176 |
| 1170 | 85.2 | 1.1 | 107 | 1 | CQ074435 | ACCESSION:CQ074435 |
| 1171 | 85.2 | 1.1 | 107 | 1 | CQ105333 | ACCESSION:CQ105333 |
| 1172 | 85.2 | 1.1 | 107 | 1 | CQ144027 | ACCESSION:CQ144027 |
| 1173 | 85.2 | 1.1 | 107 | 1 | CQ179507 | ACCESSION:CQ179507 |
| 1174 | 85.2 | 1.1 | 107 | 1 | CQ203866 | ACCESSION:CQ203866 |
| 1175 | 85.2 | 1.1 | 107 | 1 | CQ227234 | ACCESSION:CQ227234 |
| 1176 | 85.2 | 1.1 | 107 | 1 | CQ265383 | ACCESSION:CQ265383 |
| 1177 | 85.2 | 1.1 | 107 | 1 | CQ302469 | ACCESSION:CQ302469 |
| 1178 | 85.2 | 1.1 | 107 | 1 | CQ339647 | ACCESSION:CQ339647 |
| c1179 | 82.6 | 1.1 | 108 | 1 | CQ751265 | ACCESSION:CQ751265 |
| 1180 | 81.1 | 1.1 | 100 | 1 | HS97F2F | ACCESSION:Z64129 |
| 1181 | 79.2 | 1.1 | 111 | 1 | A46018 | ACCESSION:A46018 |
| 1182 | 79.2 | 1.1 | 111 | 1 | A57054 | ACCESSION:A57054 |
| 1183 | 79.2 | 1.1 | 111 | 1 | A60119 | ACCESSION:A60119 |
| 1184 | 79.2 | 1.1 | 111 | 1 | A79469 | ACCESSION:A79469 |
| 1185 | 79.2 | 1.1 | 111 | 1 | A80203 | ACCESSION:A80203 |
| 1186 | 79.2 | 1.1 | 111 | 1 | A80312 | ACCESSION:A80312 |
| 1187 | 79.2 | 1.1 | 111 | 1 | AR035355 | ACCESSION:AR035355 |
| 1188 | 79.2 | 1.1 | 111 | 1 | AR036196 | ACCESSION:AR036196 |
| 1189 | 79.2 | 1.1 | 111 | 1 | AR036716 | ACCESSION:AR036716 |
| 1190 | 79.2 | 1.1 | 111 | 1 | AR076978 | ACCESSION:AR076978 |
| 1191 | 79.2 | 1.1 | 111 | 1 | AR094699 | ACCESSION:AR094699 |
| 1192 | 79.2 | 1.1 | 111 | 1 | AR183934 | ACCESSION:AR183934 |
| 1193 | 79.2 | 1.1 | 111 | 1 | AR343146 | ACCESSION:AR343146 |
| 1194 | 79.2 | 1.1 | 111 | 1 | AR344303 | ACCESSION:AR344303 |

ALIGNMENTS

RESULT 1
AX355872

```
LOCUS       AX355872               2930 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0204678.
ACCESSION   AX355872
VERSION     AX355872.1  GI:18620523
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Keith,J.C., Mccoy,J.M. and Mi,S.
  TITLE     Methods and compositions for diagnosing and treating preeclampsia
            and gestational trophoblast disorders
  JOURNAL   Patent: WO 0204678-A 1 17-JAN-2002;
            GENETICS INSTITUTE, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .2930
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             930. .2546
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD23000.1"
                     /db_xref="GI:18620524"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"

  Query Match           38.6%;  Score 2890.1;  DB 1;  Length 2930;
  Best Local Similarity  98.3%;  Pred. No. 0;
  Matches 2873;  Conservative   48;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          8 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 67

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 127

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         128 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 187

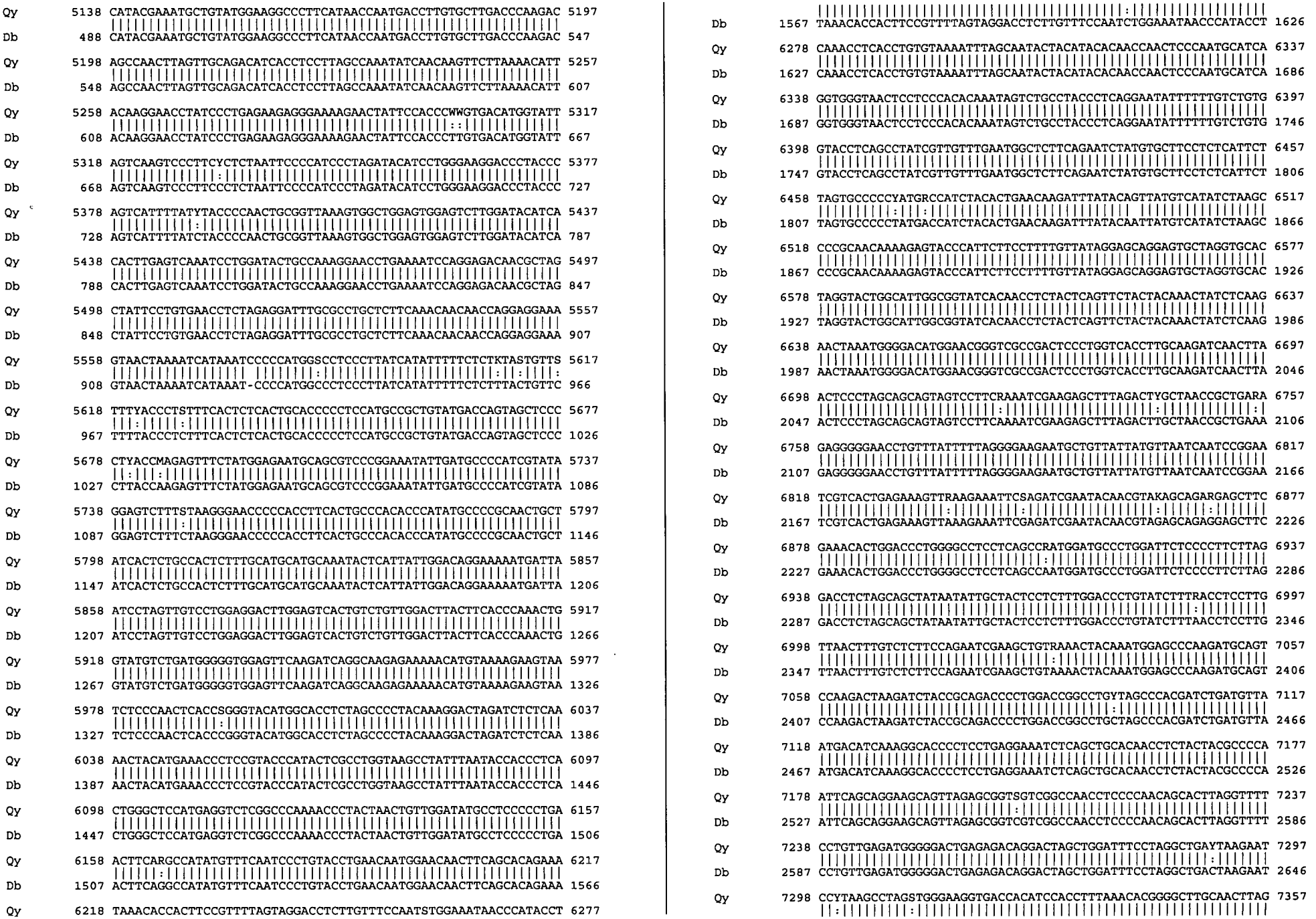
Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         188 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 247

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         248 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 307

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         308 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 367

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         368 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 427

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         428 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 487
```



```
Db       2647 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2706

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2707 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2766

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
             |||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db       2767 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2826

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2886

Qy       7538 GCTCTGTTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             |||||||||||||||||||||||||||||||||||||:||||
Db       2887 GCTCTGTTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 2930


RESULT 2
AF208161
LOCUS       AF208161               2930 bp    mRNA    linear   PRI 22-FEB-2000
DEFINITION  Homo sapiens syncytin precursor, mRNA, complete cds.
ACCESSION   AF208161
VERSION     AF208161.1  GI:6760400
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2930)
  AUTHORS   Mi,S., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L.,
            LaVallie,E., Tang,X.Y., Edouard,P., Howes,S., Keith,J.C. Jr. and
            McCoy,J.M.
  TITLE     Syncytin is a captive retroviral envelope protein involved in human
            placental morphogenesis
  JOURNAL   Nature 403 (6771), 785-789 (2000)
   PUBMED   10693809
REFERENCE   2  (bases 1 to 2930)
  AUTHORS   Sha,M., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L.,
            LaVallie,E., Tang,X., Edouard,P., Howes,S., Keith,J.C. Jr. and
            McCoy,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-1999) Genetics Institute, 87 Cambridge Park
            Drive, Cambridge, MA 02140, USA
FEATURES             Location/Qualifiers
     source          1. .2930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="AJ172-2s"
                     /tissue_type="testis"
     repeat_region   1. .2930
                     /rpt_family="Human endogeneous retrovirus HERV-W"
     5'UTR           1. .929
     CDS             930. .2546
                     /note="envelope protein"
                     /codon_start=1
                     /product="syncytin precursor"
                     /protein_id="AAF28334.1"
                     /db_xref="GI:6760401"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     sig_peptide     930. .989
     mat_peptide     990. .2543
                     /product="syncytin"
     misc_feature    1869. .1880
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="furin; cleavage site"
     misc_feature    2043. .2120
                     /note="Region: immunosuppresive region"
     misc_feature    2235. .2357
                     /note="transmembrane-region site"
     3'UTR           2547. .2930

  Query Match             38.6%;  Score 2890.1;  DB 1;  Length 2930;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2873;  Conservative   48;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:|||||||||||||||||||:||||||:||||||||||||||||:|:||:||
Db          8 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 67

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 127

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        128 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 187

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 247

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 307

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 367

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 427

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 487

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 547

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 607

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             ||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db        608 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 667

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 727

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 787

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 847

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
```

```
Db        848 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 907
Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
Db        908 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 966
Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
Db        967 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1026
Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
Db       1027 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1086
Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
Db       1087 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1146
Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
Db       1147 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1206
Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
Db       1207 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1266
Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
Db       1267 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1326
Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
Db       1327 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1386
Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
Db       1387 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1446
Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
Db       1447 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1506
Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
Db       1507 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1566
Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
Db       1567 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1626
Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
Db       1627 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1686
Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
Db       1687 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1746
Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
Db       1747 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1806
Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
Db       1807 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1866
Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
Db       1867 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1926
Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
Db       1927 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1986
Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
Db       1987 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2046
Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
Db       2047 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2106
Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
Db       2107 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2166
Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
Db       2167 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2226
Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
Db       2227 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2286
Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
Db       2287 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2346
Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
Db       2347 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2406
Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
Db       2407 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2466
Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
Db       2467 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2526
Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
Db       2527 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2586
Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
Db       2587 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2646
Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
Db       2647 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2706
Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
Db       2707 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2766
Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
Db       2767 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2826
Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
Db       2827 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2886
Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
Db       2887 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 2930

RESULT 3
AR177269
LOCUS       AR177269               2946 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6312921.
ACCESSION   AR177269
VERSION     AR177269.1  GI:17919624
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2946)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Evans,C.,
            Merberg,D., Mi,S. and Treacy,M.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 6312921-A 3 06-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .2946
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             |||||||:|||||||||||||||||||||:|||||||:|||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545

Qy       5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605

Qy       5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
             ||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||
Db        606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665

Qy       5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
             |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725

Qy       5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
             ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785

Qy       5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845

Qy       5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905

Qy       5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
             |||||||||||||||| ||||||||:|||||||||||||||||||||:||:||||:
Db         906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964

Qy       5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
             |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024

Qy       5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
             ||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084

Qy       5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144

Qy       5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204

Qy       5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264

Qy       5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324

Qy       5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384

Qy       6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444

Qy       6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504

Qy       6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564

Qy       6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624

Qy       6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684

Qy       6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744

Qy       6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804

Qy       6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
             ||||||||||:|||:|||||||||||||||||||||||||||||| ||||||||||||||||
Db        1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864

Qy       6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924
```

```
Qy       6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984

Qy       6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044

Qy       6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
             |||||||||||||||||||||||:|||||||||||||||||||:||||||||||||:|
Db       2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104

Qy       6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164

Qy       6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
             |||||||||||||||||:|||||||||:||||||||||||||||:||||||:|||||||
Db       2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224

Qy       6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404

Qy       7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464

Qy       7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524

Qy       7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584

Qy       7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy       7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
             ||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy       7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy       7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
             ||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||||
Db       2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy       7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy       7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||||||||||:||||:
Db       2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929


RESULT 4
BD086133
LOCUS       BD086133               2946 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Secreted proteins and polynucleotides encoding them.
ACCESSION   BD086133
VERSION     BD086133.1  GI:22631743
KEYWORDS    JP 2001524490-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2946)
  AUTHORS   Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C.,
            Merberg,D. and Treacy,M.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: JP 2001524490-A 2 04-DEC-2001;
            GENETICS INSTITUTE INC
COMMENT     OS   Unidentified
            PN   JP 2001524490-A/2
            PD   04-DEC-2001
            PF   17-NOV-1998 JP 2000522128
            PR   21-NOV-1997 US    08/976110,18-MAY-1998 US    09/080478 PR
            20-OCT-1998 US    09/175928
            PI   KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A COLLINS
            PI   RACIE,
            PI   CHERYL EVANS,DAVID MERBERG,MAURICE TREACY
            PC   C07K14/47,C12N5/10,C12N15/09,C12P21/02,C12N5/00,C12N15/00 CC
            Strandedness: Double;
            CC   Topology: Linear;
            CC   Secreted proteins and polynucleotides encoding them FH   Key
                      Location/Qualifiers
            FT   source          1. .2946
            FT                   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .2946
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"

  Query Match             38.6%;  Score 2889.1;  DB 1;  Length 2946;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2872;  Conservative   49;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             |||||||||||||||||||||||:||||||:||||||||||||||||||:|:||:||
Db          6 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGAC 65

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 125

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             ||:|||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db         126 TAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 185

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 245

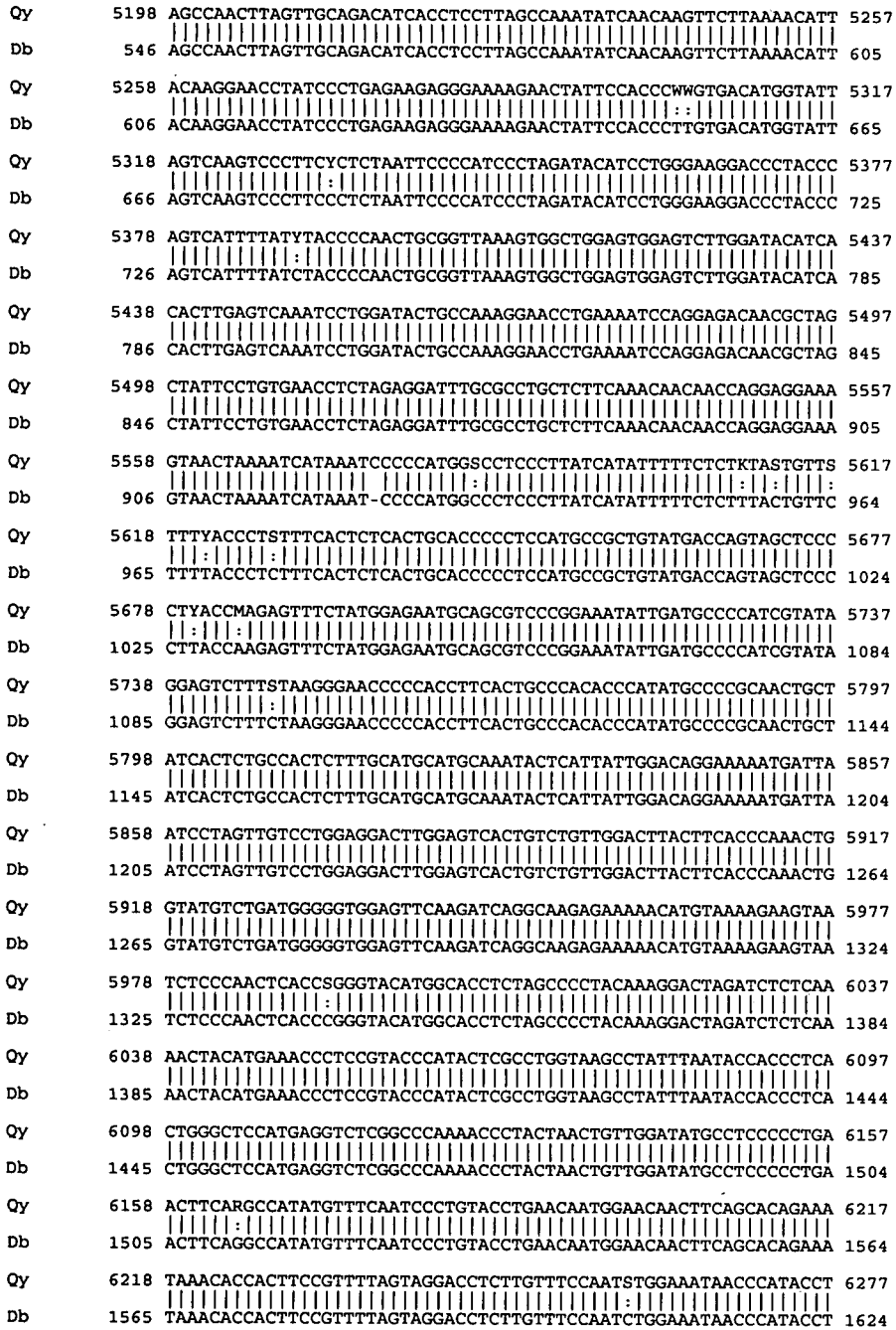
Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 4957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTA 305

Qy       4958 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 5017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 TACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGA 365

Qy       5018 ATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 5077
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 ATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTG 425

Qy       5078 YTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 5137
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 CTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCC 485

Qy       5138 CATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 5197
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         486 CATACGAAATGCTGTATGGAAGCCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGAC 545
```


Qy 5198 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 5257
Db 546 AGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATT 605
Qy 5258 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATT 5317
Db 606 ACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATT 665
Qy 5318 AGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 5377
Db 666 AGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCC 725
Qy 5378 AGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 5437
Db 726 AGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA 785
Qy 5438 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 5497
Db 786 CACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAG 845
Qy 5498 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
Db 846 CTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 905
Qy 5558 GTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTS 5617
Db 906 GTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTC 964
Qy 5618 TTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 5677
Db 965 TTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCC 1024
Qy 5678 CTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 5737
Db 1025 CTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATA 1084
Qy 5738 GGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 5797
Db 1085 GGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCT 1144
Qy 5798 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 5857
Db 1145 ATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTA 1204
Qy 5858 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 5917
Db 1205 ATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTG 1264
Qy 5918 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 5977
Db 1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324
Qy 5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
Db 1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384
Qy 6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
Db 1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444
Qy 6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
Db 1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504
Qy 6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
Db 1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564
Qy 6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
Db 1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624
Qy 6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
Db 1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684
Qy 6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
Db 1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744
Qy 6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
Db 1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804
Qy 6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
Db 1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864
Qy 6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
Db 1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924
Qy 6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
Db 1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984
Qy 6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
Db 1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044
Qy 6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
Db 2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104
Qy 6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
Db 2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164
Qy 6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
Db 2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224
Qy 6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
Db 2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284
Qy 6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
Db 2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344
Qy 6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
Db 2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404
Qy 7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
Db 2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464
Qy 7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
Db 2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524
Qy 7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
Db 2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584
Qy 7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
Db 2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644
Qy 7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
Db 2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704
Qy 7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417

```
          :|||||||
```

```
Db      1265 GTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAA 1324

Qy      5978 TCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 6037
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db      1325 TCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAA 1384

Qy      6038 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 6097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1385 AACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCA 1444

Qy      6098 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 6157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1445 CTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGA 1504

Qy      6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 6217
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1505 ACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAA 1564

Qy      6218 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCT 6277
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db      1565 TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCT 1624

Qy      6278 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 6337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1625 CAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCA 1684

Qy      6338 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 6397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1685 GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTG 1744

Qy      6398 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 6457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1745 GTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCT 1804

Qy      6458 TAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGC 6517
             ||||||||||:|||:|||||||||||||||||||||||||||| ||||||||||||||||||
Db      1805 TAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAAGC 1864

Qy      6518 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 6577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1865 CCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC 1924

Qy      6578 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 6637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1925 TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAG 1984

Qy      6638 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 6697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1985 AACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTA 2044

Qy      6698 ACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARA 6757
             ||||||||||||||||||||||||:|||||||||||||||||||:||||||||||||:|
Db      2045 ACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAA 2104

Qy      6758 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 6817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2105 GAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAA 2164

Qy      6818 TCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC 6877
             ||||||||||||||||:||||||||:||||||||||||||||:|||||:|||||||
Db      2165 TCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTC 2224

Qy      6878 GAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAG 6937
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      2225 GAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAG 2284

Qy      6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db      2285 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTG 2344

Qy      6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGT 7057
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      2345 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGT 2404

Qy      7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTA 7117
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db      2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTA 2464

Qy      7118 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 7177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2465 ATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCA 2524

Qy      7178 ATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 7237
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db      2525 ATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTT 2584

Qy      7238 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAAT 7297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db      2585 CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAT 2644

Qy      7298 CCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 7357
             ||:||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2645 CCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAG 2704

Qy      7358 YTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAA 7417
             :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      2705 CTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAA 2764

Qy      7418 AGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 7477
             ||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||||
Db      2765 AGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATA 2824

Qy      7478 AACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 7537
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2825 AACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGA 2884

Qy      7538 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||||||||||||||:||||:
Db      2885 GCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2929
```

```
RESULT 6
BD196236
LOCUS       BD196236               2938 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196236
VERSION     BD196236.1  GI:33006006
KEYWORDS    JP 2002512530-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 2 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/2
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
           MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
           Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC           diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2938
            FT                   /organism='Unidentified'.
```

```
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match           36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              |||||||||||||:||||::|||||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
              ||||||||||:|||||||||||||||:||||||||||||||||||||||||||||||||||
Db        540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy       1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

Qy       1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy       1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
                ||||||||||||| | ||      |       |||    |||||||:||||||||||||
Db        718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy       1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy       1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
              |||||||||||||||:|||||||||||||||||:|||||:|||||:||||||||||||||
Db        833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy       1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952

Qy       1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012

Qy       1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072

Qy       1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132

Qy       1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192

Qy       1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252

Qy       1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552

Qy       2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611

Qy       2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
              ||||||| |||||||||||||||||||||||||| ||:|||||||||||||||| |||||
Db        1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668

Qy       2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728

Qy       2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788

Qy       2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847

Qy       2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906

Qy       2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966

Qy       2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026
```

```
Qy       2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086

Qy       2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db       2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||||:|||||||
Db       2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             |||||||||||||||||||||||||||||||:||:|::|||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:||||||||||:||||||||||||:||||||||||:|:|:|||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||||||    ||||||||:||||:||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:|||||||:::|||||||||||||||||||||||||||||||||||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC-------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:||||||||:||||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:||||||:|||||||||||||||:|||:||||||:||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 7
BD267478
LOCUS       BD267478               2938 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267478
VERSION     BD267478.1  GI:33077246
KEYWORDS    JP 2002534980-A/19.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 19 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/19
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .2938
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match           36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

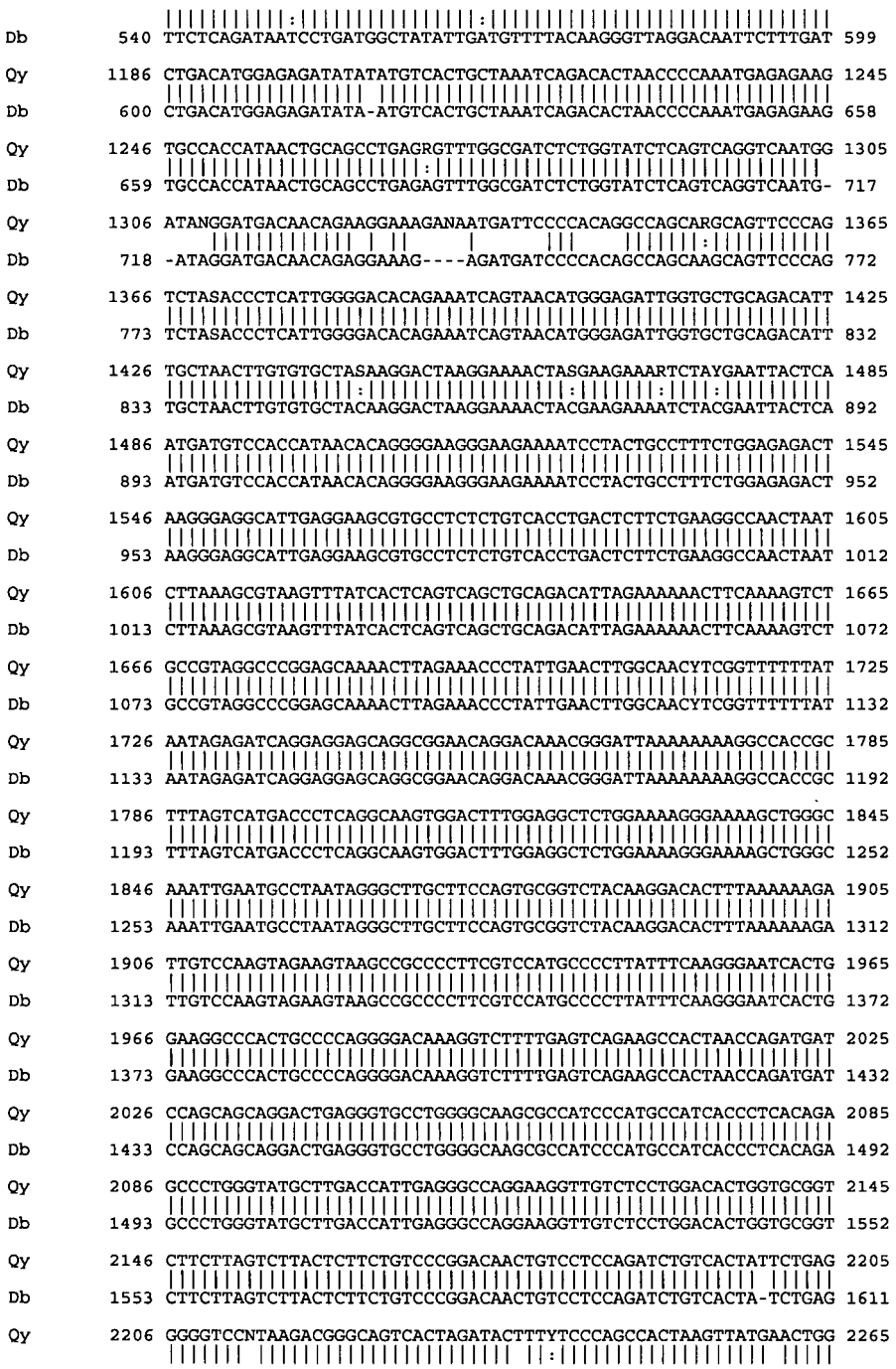
Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              |||||||||||||:|||||::|||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
```


Db 540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599
Qy 1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
Db 600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658
Qy 1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
Db 659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717
Qy 1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
Db 718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772
Qy 1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
Db 773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832
Qy 1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
Db 833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892
Qy 1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
Db 893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952
Qy 1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
Db 953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012
Qy 1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
Db 1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072
Qy 1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
Db 1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132
Qy 1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
Db 1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192
Qy 1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
Db 1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252
Qy 1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
Db 1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312
Qy 1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
Db 1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372
Qy 1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
Db 1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432
Qy 2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
Db 1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492
Qy 2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
Db 1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552
Qy 2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
Db 1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611
Qy 2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
Db 1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668
Qy 2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
Db 1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728
Qy 2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
Db 1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788
Qy 2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
Db 1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847
Qy 2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
Db 1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906
Qy 2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
Db 1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966
Qy 2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
Db 1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026
Qy 2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
Db 2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086
Qy 2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
Db 2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146
Qy 2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
Db 2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206
Qy 2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
Db 2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266
Qy 2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
Db 2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326
Qy 2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
Db 2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384
Qy 2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
Db 2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444
Qy 3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
Db 2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504
Qy 3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
Db 2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560
Qy 3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
Db 2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620
Qy 3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
Db 2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679
Qy 3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
Db 2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

```
Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
              |||||||:|:||||||:::|||||||||||||||||||                |||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC-------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
              :|||||||||||||:||||||||:|||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
              |:||:||||||:|||||||||||||||:|||:||||||:||||||||||||::|||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 8
AX000957
LOCUS       AX000957               2938 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 2 from Patent WO9902696.
ACCESSION   AX000957
VERSION     AX000957.1  GI:7241199
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 2 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:||||||||:||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              ||||||||||||:||||::|||||||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
              ||||||||||:|||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy       1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

Qy       1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
              |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy       1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
              |||||||||||| | ||   |   |||   |||||||:||||||||||
Db        718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy       1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy       1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
              |||||||||||||||||:|||||||||||||||:|||||||:||||:||||||||||||
Db        833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy       1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952

Qy       1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012

Qy       1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072

Qy       1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132

Qy       1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192

Qy       1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252

Qy       1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552

Qy       2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611

Qy       2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
             |||||||| |||||||||||||||||||||||| ||:|||||||||||||||||||| |||||
Db       1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668

Qy       2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728

Qy       2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788

Qy       2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847

Qy       2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906

Qy       2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966

Qy       2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026

Qy       2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086

Qy       2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db       2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||:|||||||||
Db       2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             |||||||||||||||||||||||||||||||||:||:|::||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:||||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:|||||||||||:|||||||||||||:||||||||||:|:|:|||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||||    ||||||||:||||:|||||||||||||||||:|||||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||||:|||||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||                ||||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:|||||||||:|||||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:|||||||:|||||||||||||||:|||:|||||:||||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 9
AX027471
LOCUS       AX027471               2938 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 21 from Patent FR2788784.
ACCESSION   AX027471
VERSION     AX027471.1  GI:10188435
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 21 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

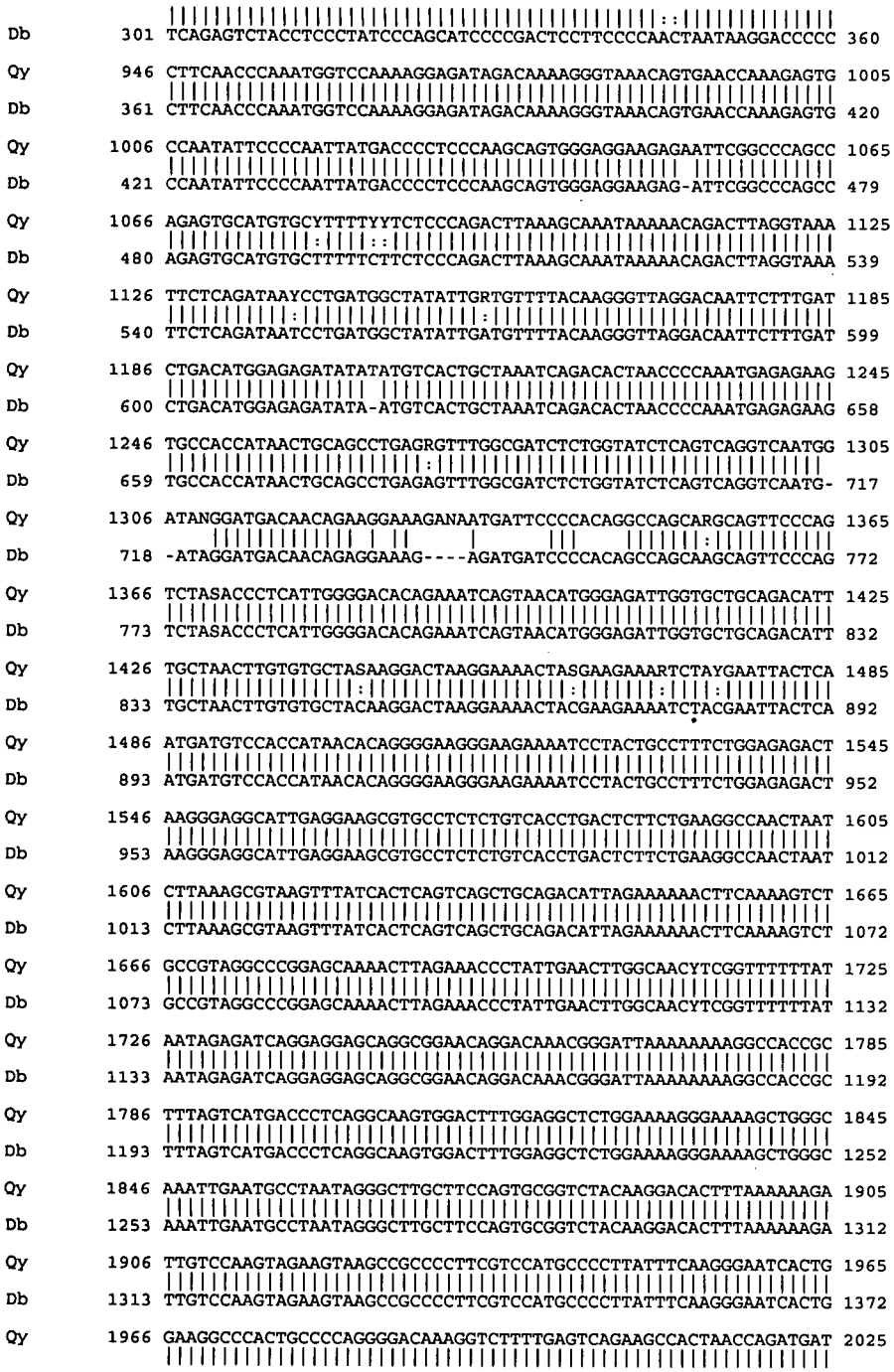
Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:|||||||:|||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
```


Db 301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360
Qy 946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
Db 361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420
Qy 1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
Db 421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479
Qy 1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
Db 480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539
Qy 1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
Db 540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599
Qy 1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
Db 600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658
Qy 1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
Db 659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717
Qy 1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
Db 718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772
Qy 1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
Db 773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832
Qy 1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
Db 833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892
Qy 1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
Db 893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952
Qy 1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
Db 953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012
Qy 1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
Db 1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072
Qy 1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
Db 1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132
Qy 1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
Db 1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192
Qy 1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
Db 1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252
Qy 1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
Db 1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312
Qy 1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
Db 1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372
Qy 1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
Db 1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432
Qy 2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
Db 1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492
Qy 2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
Db 1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552
Qy 2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
Db 1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611
Qy 2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
Db 1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668
Qy 2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
Db 1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728
Qy 2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
Db 1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788
Qy 2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
Db 1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847
Qy 2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
Db 1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906
Qy 2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
Db 1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966
Qy 2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
Db 1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026
Qy 2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
Db 2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086
Qy 2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
Db 2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146
Qy 2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
Db 2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206
Qy 2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
Db 2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266
Qy 2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
Db 2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326
Qy 2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
Db 2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384
Qy 2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
Db 2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444
Qy 3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
Db 2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

```
Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||||||    ||||||||:||||:||||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::|||||||||||||||||||||||||||||:||||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||                ||||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:||||||||:|||||||||||||:||:|||::|||||| ||||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:||||||:|||||||||||||||:|||:|||||:||||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:||||||||:||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 10
AF072499
LOCUS       AF072499                2938 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
ACCESSION   AF072499
VERSION     AF072499.1  GI:4262283
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 2938)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.6A1"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     misc_feature    1. .2938
                     /note="corresponds to gag and pol"
     repeat_region   1. .2938
                     /rpt_family="HERV-W"

  Query Match             36.1%;  Score 2706.61;  DB 1;  Length 2938;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 2831;  Conservative   90;  Mismatches   17;  Indels   36;  Gaps   15;

Qy        586 CAMSRACGGACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 645
              ||:::|||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACGACGGACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGA 60

Qy        646 CGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 705
              ||||||||||:|||||||:|||||||||||||||||||||||||||||||||||||||||
Db         61 CGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATA 120

Qy        706 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA 180

Qy        766 GCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 825
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTAGACTTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTT 240

Qy        826 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCT 300

Qy        886 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCC 945
              ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db        301 TCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCC 360

Qy        946 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTG 420

Qy       1006 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCC 1065
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        421 CCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG-ATTCGGCCCAGCC 479

Qy       1066 AGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 1125
              |||||||||||||:||||::|||||||||||||||||||||||||||||||||||||||
Db        480 AGAGTGCATGTGCTTTTTCTTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAA 539

Qy       1126 TTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGAT 1185
              ||||||||||:|||||||||||||||||:|||||||||||||||||||||||||||||||
Db        540 TTCTCAGATAATCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGAT 599

Qy       1186 CTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 1245
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        600 CTGACATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAG 658

Qy       1246 TGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGG 1305
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        659 TGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG- 717

Qy       1306 ATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAG 1365
              |||||||||||||| | ||     |      |||    |||||||:||||||||||||
Db        718 -ATAGGATGACAACAGAGGAAAG----AGATGATCCCCACAGCCAGCAAGCAGTTCCCAG 772

Qy       1366 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATT 832

Qy       1426 TGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCA 1485
              |||||||||||||||||:|||||||||||||||:|||||||:||||:||||||||||||
Db        833 TGCTAACTTGTGTGCTACAAGGACTAAGGAAAACTACGAAGAAAATCTACGAATTACTCA 892

Qy       1486 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACT 952

Qy       1546 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 AAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAAT 1012

Qy       1606 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1013 CTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCT 1072

Qy       1666 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTAT 1132

Qy       1726 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGC 1192

Qy       1786 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGC 1252

Qy       1846 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 AAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGA 1312

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1372

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 1432

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 1492

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 2145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGT 1552

Qy       2146 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAG 2205
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1553 CTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTA-TCTGAG 1611

Qy       2206 GGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGG 2265
             ||||||| |||||||||||||||||||||| ||:|||||||||||||||||| |||||
Db       1612 GGGGTCC-TAAGACGGGCAGTCACTAGATAC-TTCTCCCAGCCACTAAGTTATG-ACTGG 1668

Qy       2266 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 2325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTA 1728

Qy       2326 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 2385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1729 GGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACA 1788

Qy       2386 CCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 2445
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 CCCGTTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGG 1847

Qy       2446 ACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 2505
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       1848 ACAATATGGACGAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTT 1906

Qy       2506 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 CCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGA 1966

Qy       2566 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 TTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATT 2026

Qy       2626 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCT 2086

Qy       2686 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTG 2146

Qy       2746 TGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCT 2805
             ||::|||||||||||||||||||||||||||||||||||:|||||||||||:|||||||||
Db       2147 TGCTTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACACTCCTGGACCTTAAGGATGCCT 2206

Qy       2806 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAA 2266

Qy       2866 ACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATC 2925
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||||:|||||||
Db       2267 ACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATC 2326

Qy       2926 TATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCC 2985
             |||||||||||||||||||||||||||||||:||:|::|||||||||||||  |||||||
Db       2327 TATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCC 2384

Qy       2986 TTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCA 3045
             |||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       2385 TTCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCA 2444

Qy       3046 CCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTC 3105
             |||||||:|||||:|||||||||||:||||||||||||:|||||||||||:|:|:||||
Db       2445 CCCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTC 2504

Qy       3106 ARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCC 3165
             |:|||||||||||||||    ||||||||:||||:||||||||||||||||:|||||
Db       2505 AACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCC 2560

Qy       3166 TCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAAC 3225
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db       2561 TCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAAC 2620

Qy       3226 TAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAA 3285
             ||||:|::|||||||||:|||:|||:||||||||||:|||||| |||||||:||||:||
Db       2621 TAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTTTGGCGAA 2679

Qy       3286 ATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTA 3345
             |||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||:||
Db       2680 ATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTA 2739

Qy       3346 AGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAG 3405
             |||||||:|:||||||:::|||||||||||||||||||                |||||||
Db       2740 AGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC-------------TAACCCAAG 2785

Qy       3406 YCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACA 3465
             :|||||||||||:|||||||:|||||||||||||:||:|||::|||||| |||||||
Db       2786 CCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGTCACAG-AAAAAACA 2844

Qy       3466 GRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYAC 3525
             |:||:||||||:|||||||||||||||:|||:|||||:||||||||||::||||::||
Db       2845 GGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCGTAC 2904

Qy       3526 CTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||:|||||||:||||||||||||||||||||||
Db       2905 CTGACTAAGGAAATTGATGTAGTGGCAAAGGGTT 2938


RESULT 11
BD196243
LOCUS       BD196243               3372 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196243
VERSION     BD196243.1  GI:33006013
KEYWORDS    JP 2002512530-A/9.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 3372)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
```

```
  JOURNAL   Patent: JP 2002512530-A 9 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/9
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .3372
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .3372
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
Db            61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
Db           121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
Db           181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
Db           239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
Db           299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
Db           359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
Db           415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
Db           475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
Db           534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
Db           594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
Db           640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
Db           699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
Db           759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
Db           819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
Db           879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
Db           939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
Db           999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
Db           1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
Db           1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
Db           1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
Db           1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
Db           1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356

Qy       4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
Db           1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416

Qy       4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
Db           1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476

Qy       4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
Db           1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536

Qy       4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
Db           1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596

Qy       4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
Db           1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656

Qy       4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
Db           1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716

Qy       4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
```

```
Db      1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776

Qy      4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836

Qy      4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:|||:|||||||||||||:|||||||:|:||||||||||||||||:||
Db      1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896

Qy      4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
             ||||||||||||||||||:||||||:||||||||||||||||:|:||:||||||||||||
Db      1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956

Qy      4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db      1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016

Qy      4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db      2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076

Qy      4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136

Qy      4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196

Qy      4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db      2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256

Qy      5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db      2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316

Qy      5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376

Qy      5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436

Qy      5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496

Qy      5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
             ||||||||||||||||||   |||  | ||||||||||::||||||||||||||||||||
Db      2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554

Qy      5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

Qy      5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy      5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy      5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785

RESULT 12
```

```
BD267485
LOCUS       BD267485               3372 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267485
VERSION     BD267485.1  GI:33077253
KEYWORDS    JP 2002534980-A/26.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3372)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 26 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/26
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .3372
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .3372
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match           35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative   141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy      2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db         1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy      2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy      2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:|||||||||||||||||||||||:||||||||
Db       121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

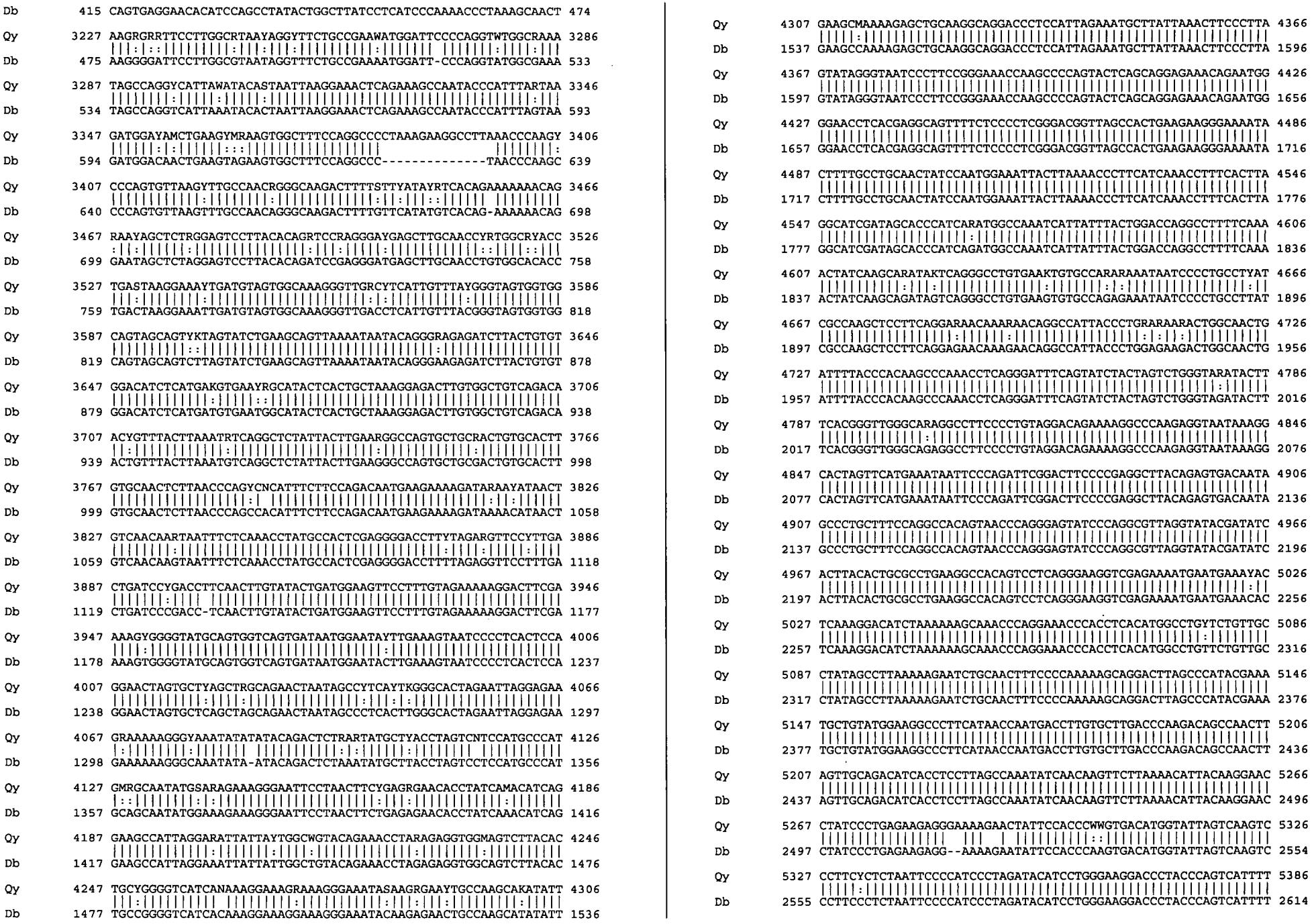
Qy      2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||:||:|::|||||||||||  ||||||||
Db       181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy      2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||||:||:|||::|||||||||||||||||||||||||||
Db       239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy      3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:||||||||||:||||||||||||:||||||||||:|:|:|||||
Db       299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy      3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||    |||||||||:||||:|||||||||||||||:|||||||
Db       359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy      3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
```


Db 415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474
Qy 3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
Db 475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533
Qy 3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
Db 534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593
Qy 3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
Db 594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC----------TAACCCAAGC 639
Qy 3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
Db 640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698
Qy 3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
Db 699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758
Qy 3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
Db 759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818
Qy 3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
Db 819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878
Qy 3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
Db 879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938
Qy 3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
Db 939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998
Qy 3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
Db 999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058
Qy 3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
Db 1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118
Qy 3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
Db 1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177
Qy 3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
Db 1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237
Qy 4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
Db 1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297
Qy 4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
Db 1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356
Qy 4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
Db 1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416
Qy 4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
Db 1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476
Qy 4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
Db 1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536
Qy 4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
Db 1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596
Qy 4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
Db 1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656
Qy 4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
Db 1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716
Qy 4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
Db 1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776
Qy 4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
Db 1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836
Qy 4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
Db 1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896
Qy 4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
Db 1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956
Qy 4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
Db 1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016
Qy 4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
Db 2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076
Qy 4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
Db 2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136
Qy 4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
Db 2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196
Qy 4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
Db 2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256
Qy 5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
Db 2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316
Qy 5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
Db 2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376
Qy 5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
Db 2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436
Qy 5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
Db 2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496
Qy 5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
Db 2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554
Qy 5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
Db 2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

```
Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785


RESULT 13
AX000964
LOCUS       AX000964               3372 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 9 from Patent WO9902696.
ACCESSION   AX000964
VERSION     AX000964.1  GI:7241206
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 3372)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 9 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .3372
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||||:||||||||||:|||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||:|||||||||||||||||||||||:||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||:||:|::|||||||  ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||:||:|||::|||||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:|||||||||||:|||||||||||||:|||||||||||:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||    |||||||||:|||||:|||||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             |||||||||:::|||||||||||||||||||||||||||||:||||||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::||||||||||:|||:|||:||||||||||:|||||||| ||||||||:||||:|||

Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db        534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
             ||||||:|:||||||:::|||||||||||||||||||||||||||||||||||||||||:
Db        594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
             ||||||||||||:||||||||:|||||||||||||||:||:|||::|||||| ||||||||
Db        640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
             :||:|||||||:|||||||||||||||||:|||:|||||:||||||||||||::|||||:|||
Db        699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
             |||:||||||||:|||||||||||||||||||||:|:||||||||||:||||||||||||
Db        759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
             ||||||||||::|||||||||||||||||||||||||||||:|||||||||||||||||||
Db        819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
             |||||||||||||:|||||::|||||||||||||||||||||||||||||||||||||||
Db        879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
             ||:|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
             |||||||||||||||||:| ||||||||||||||||||||||||||:||:||||||
Db        999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||||:|||||||||||||||||||||||||||||||||||||:||||:|||||:||||
Db       1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
             ||||:|||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
             ||||||||||||:||||:|||||||||||||||:|||:|:|||||||||||||||||||||
Db       1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
             |:||||||||:|||||| ||||||||||:|||||:|||||||||| ||||||||||||||
Db       1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356

Qy       4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
             |::|||||||:|:|||||||||||||||||||||:|||:||||||||||||:|||||||||
Db       1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416

Qy       4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
             ||||||||||||:|||||||:||||:|||||||||||||:||||||||:||||||||||||
Db       1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476

Qy       4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
             |||:|||||||| |||||||||:|||||||||||||:|||:|||:|||||||||||||||
Db       1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536

Qy       4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596
```

```
Qy       4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656

Qy       4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716

Qy       4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:|||:|||||||||||||||:|||||||:|:||||||||||||||||:||
Db       1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
             ||||||||||||||||||:||||||:|||||||||||||||||:|:||:|||||||||||
Db       1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956

Qy       4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016

Qy       4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076

Qy       4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136

Qy       4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196

Qy       4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256

Qy       5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316

Qy       5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376

Qy       5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496

Qy       5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
             ||||||||||||||||  |||  | ||||||||||::|||||||||||||||||||||||
Db       2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554

Qy       5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785
```

RESULT 14
AX027478
LOCUS AX027478 3372 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 28 from Patent FR2788784.
ACCESSION AX027478
VERSION AX027478.1 GI:10188442
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patent: FR 2788784-A 28 28-JUL-2000;
BIO MERIEUX (FR)
FEATURES Location/Qualifiers
source 1. .3372
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 35.1%; Score 2626.2; DB 1; Length 3372;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 2637; Conservative 141; Mismatches 7; Indels 26; Gaps 8;

```
Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:|||||||||||||||||||||||:|||||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||:||:|::|||||||||||  ||||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||||:||:|||::||||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

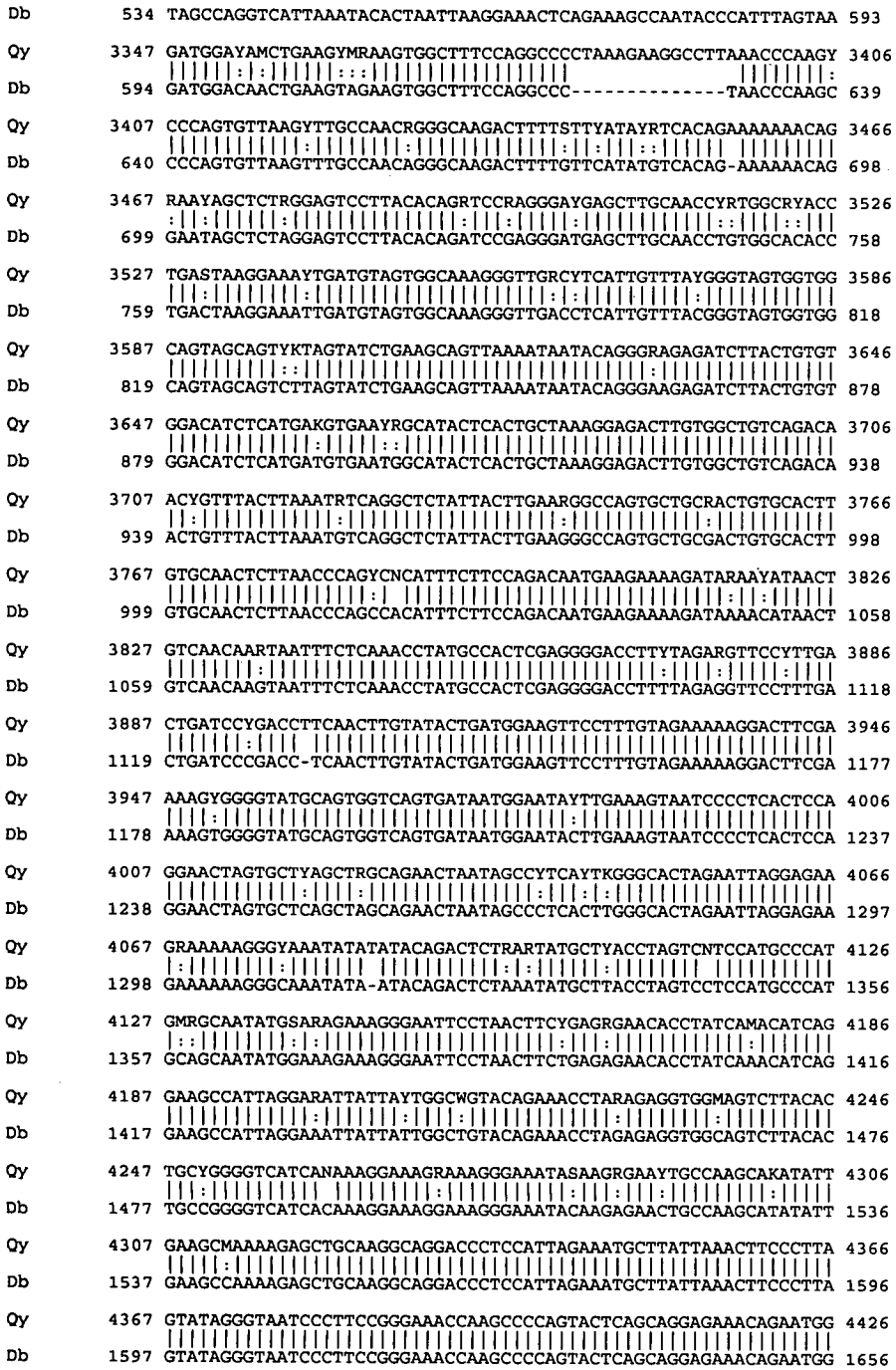
Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:||||||||||:||||||||||||:|||||||||||:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||||  ||||||||:||||:|||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||||:|||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::|||||||||:|||:|||:||||||||||:|||||| |||||||:||||:|||
Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
```


Db 534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593
Qy 3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
Db 594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC----------TAACCCAAGC 639
Qy 3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
Db 640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698
Qy 3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
Db 699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758
Qy 3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
Db 759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818
Qy 3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
Db 819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878
Qy 3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
Db 879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938
Qy 3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
Db 939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998
Qy 3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
Db 999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058
Qy 3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
Db 1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118
Qy 3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
Db 1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177
Qy 3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
Db 1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237
Qy 4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
Db 1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297
Qy 4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
Db 1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356
Qy 4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
Db 1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416
Qy 4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
Db 1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476
Qy 4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
Db 1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536
Qy 4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
Db 1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596
Qy 4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
Db 1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656
Qy 4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
Db 1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716
Qy 4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
Db 1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776
Qy 4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
Db 1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836
Qy 4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
Db 1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896
Qy 4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
Db 1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956
Qy 4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
Db 1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016
Qy 4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
Db 2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076
Qy 4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
Db 2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136
Qy 4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
Db 2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196
Qy 4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
Db 2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256
Qy 5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
Db 2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316
Qy 5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
Db 2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376
Qy 5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
Db 2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436
Qy 5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
Db 2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496
Qy 5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
Db 2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554
Qy 5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
Db 2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614
Qy 5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
Db 2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674
Qy 5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
Db 2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

```
Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785


RESULT 15
AF072502
LOCUS       AF072502               3372 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
ACCESSION   AF072502
VERSION     AF072502.1  GI:4262286
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3372)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 3372)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .3372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.Pi5T"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     misc_feature    1. .3372
                     /note="corresponds to pol"
     repeat_region   1. .3372
                     /rpt_family="HERV-W"
     misc_feature    2785. .2786
                     /note="putative splice junction"
     3'UTR           3025. .3351
                     /note="U3-R junction undetermined"
     polyA_site      3350

  Query Match             35.1%;  Score 2626.2;  DB 1;  Length 3372;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches 2637;  Conservative  141;  Mismatches    7;  Indels   26;  Gaps    8;

Qy       2747 GMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTT 2806
             |::|||||||||||||||||||||||||||||||||||:||||||||||:||||||||||
Db          1 GACTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTT 60

Qy       2807 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAA 120

Qy       2867 CCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCT 2926
             ||||:|||||||||||||||||||:||||||||||||||||||||:||||||||
Db        121 CCCAGCATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCT 180

Qy       2927 ATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCT 2986
             |||||||||||||||||||||||||||||:||:|::|||||||||||||  ||||||||
Db        181 ATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCT 238

Qy       2987 TCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCAC 3046
             ||:|||:|||||||||||||||:||:|||::|||||||||||||||||||||||||||||
Db        239 TCGGTAGGTGGATGATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCAC 298

Qy       3047 CCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCA 3106
             ||||||:|||||:||||||||||:||||||||||||:||||||||||:|:|:|||||
Db        299 CCAAGCGCTCTTCAATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCA 358

Qy       3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCT 3166
             :|||||||||||||||||    |||||||||:||||:|||||||||||||||:||||||
Db        359 ACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCT 414

Qy       3167 CAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACT 3226
             ||||||||||:::|||||||||||||||||||||||||||||:|||||||||||||||||
Db        415 CAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACT 474

Qy       3227 AAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAA 3286
             |||:|::||||||||||:|||:|||:||||||||||:|||||||  |||||||:||||:|||
Db        475 AAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-CCCAGGTATGGCGAAA 533

Qy       3287 TAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAA 3346
             ||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||:|||
Db        534 TAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAA 593

Qy       3347 GATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGY 3406
             ||||||:|:||||||:::|||||||||||||||||||  |||||||||:
Db        594 GATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------------TAACCCAAGC 639

Qy       3407 CCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAG 3466
             ||||||||||||:||||||||:||||||||||||:||:|||::|||||| |||||||||
Db        640 CCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTGTTCATATGTCACAG-AAAAAACAG 698

Qy       3467 RAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACC 3526
             :||:||||||:|||||||||||||:|||:|||||:||||||||||||::||||:|||
Db        699 GAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCACACC 758

Qy       3527 TGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGG 3586
             |||:|||||||:|||||||||||||||||||:|:||||||||:||||||||||||
Db        759 TGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGG 818

Qy       3587 CAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGT 3646
             ||||||||||::|||||||||||||||||||||||||||||||:|||||||||||||||
Db        819 CAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGT 878

Qy       3647 GGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 3706
             ||||||||||||:|||||::|||||||||||||||||||||||||||||||||||||||
Db        879 GGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACA 938

Qy       3707 ACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTT 3766
             ||:||||||||||||:|||||||||||||||||||:||||||||||||:||||||||||
Db        939 ACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTT 998

Qy       3767 GTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACT 3826
             ||||||||||||||||:| |||||||||||||||||||||||||||||:||:||||||
Db        999 GTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAACATAACT 1058

Qy       3827 GTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGA 3886
             ||||||||:|||||||||||||||||||||||||||||||||||||||:||||:|||||:||||
Db       1059 GTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGA 1118

Qy       3887 CTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 3946
             |||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGA 1177

Qy       3947 AAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCA 4006
             ||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1178 AAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCA 1237

Qy       4007 GGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAA 4066
             ||||||||||||:||||:|||||||||||||||:|||:|:|||||||||||||||||||
Db       1238 GGAACTAGTGCTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAA 1297

Qy       4067 GRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCAT 4126
             |:|||||||:|||||| ||||||||||:|:||||:|||||||| |||||||||||
Db       1298 GAAAAAAGGGCAAATATA-ATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCAT 1356

Qy       4127 GMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAG 4186
             |::|||||||:|:|||||||||||||||||||||:|||:||||||||||:|||||||
```

```
Db       1357 GCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAG 1416

Qy       4187 GAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACAC 4246
             ||||||||||||:|||||||:||||:||||||||||||:|||||||||:||||||||||
Db       1417 GAAGCCATTAGGAAATTATTATTGGCTGTACAGAAACCTAGAGAGGTGGCAGTCTTACAC 1476

Qy       4247 TGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATT 4306
             |||:||||||||||| ||||||||||:||||||||||||:|||:|||:||||||||||:|||||
Db       1477 TGCCGGGGTCATCACAAAGGAAAGGAAAGGGAAATACAAGAGAACTGCCAAGCATATATT 1536

Qy       4307 GAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 4366
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAAGCCAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTA 1596

Qy       4367 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 4426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGG 1656

Qy       4427 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 4486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GGAACCTCACGAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATA 1716

Qy       4487 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 4546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 CTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTA 1776

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1777 GGCATCGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 1836

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             |||||||||||:|||:|||||||||||||||:|||||||:|:|||||||||||||||:||
Db       1837 ACTATCAAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCCTGCCTTAT 1896

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTG 4726
             ||||||||||||||||||:||||||:|||||||||||||||||:|:||:||||||||||||
Db       1897 CGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTG 1956

Qy       4727 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTT 4786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       1957 ATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTT 2016

Qy       4787 TCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 4846
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       2017 TCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGG 2076

Qy       4847 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 4906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2077 CACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATA 2136

Qy       4907 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 4966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 GCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATC 2196

Qy       4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       2197 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 2256

Qy       5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db       2257 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTTGC 2316

Qy       5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2317 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 2376

Qy       5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2377 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 2436

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2437 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 2496

Qy       5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
             |||||||||||||||||  |||  | ||||||||::|||||||||||||||||||||||
Db       2497 CTATCCCTGAGAAGAGG--AAAAGAATATTCCACCCAAGTGACATGGTATTAGTCAAGTC 2554

Qy       5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2555 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 2614

Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 2674

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 5506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 CAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTG 2734

Qy       5507 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 5557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 TGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAA 2785
```

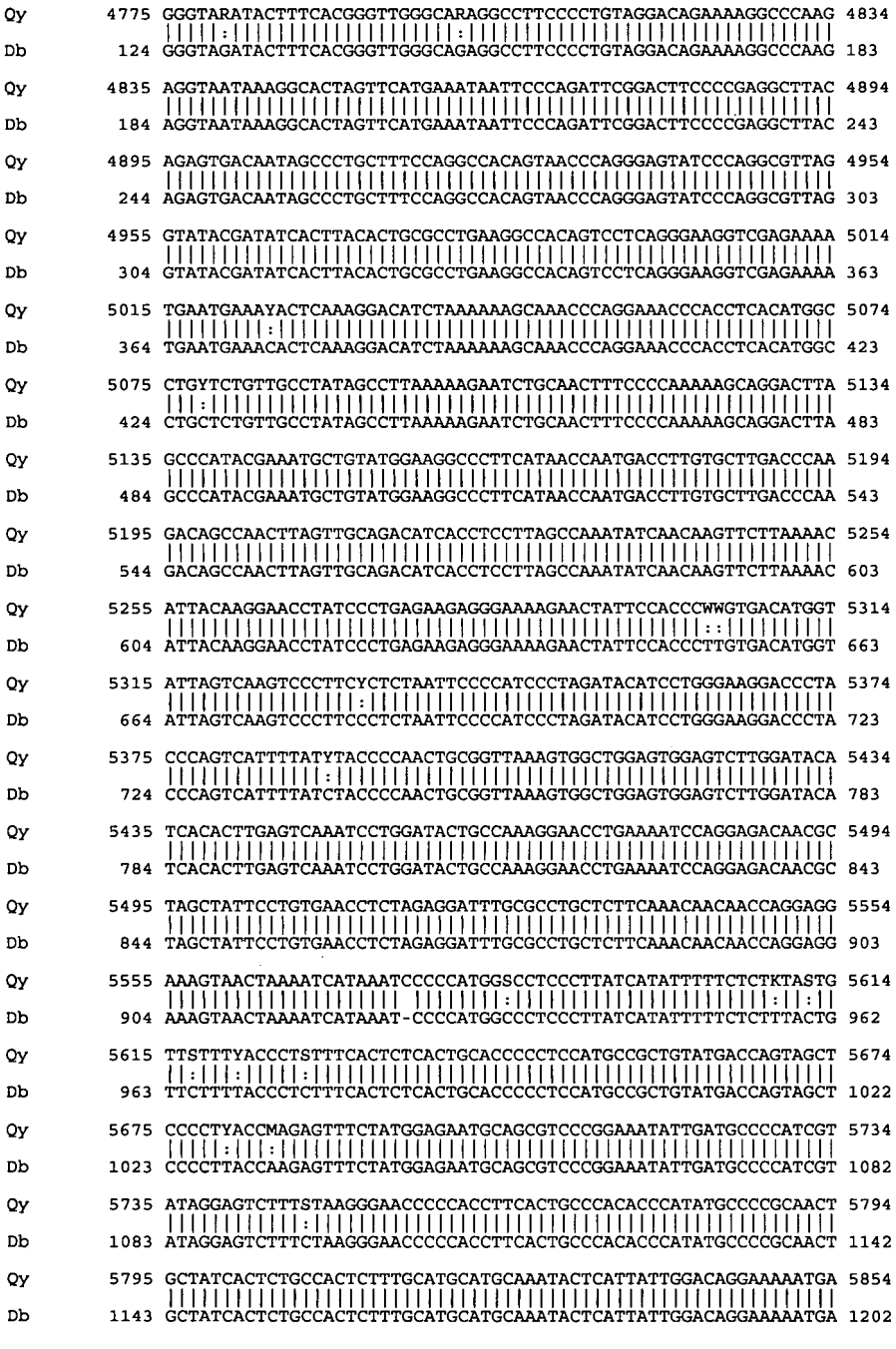
RESULT 16
BD221806

```
LOCUS       BD221806                2599 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221806
VERSION     BD221806.1  GI:33031576
KEYWORDS    JP 2002518051-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2599)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 1 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002518051-A/1
            PD   25-JUN-2002
            PF   23-JUN-1999 JP 2000556036
            PR   23-JUN-1998 FR     98/07920
            PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
            C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
            A61P25/00,
            PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
            sequence and deduced protein sequence family with  CC
            human endogenous
            CC   retroviral motifs, and their uses
            FH   Key             Location/Qualifiers
            FT   source          1. .2599
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2599
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             34.1%;  Score 2551.4;  DB 1;  Length 2599;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2552;  Conservative   39;  Mismatches    5;  Indels    2;  Gaps    2;

Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714
                |||   | | |||||||||||||||||:||||||:||||||||||||||||:|:||
Db          4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy       4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123
```


Qy 4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
Db 124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183
Qy 4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
Db 184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243
Qy 4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
Db 244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303
Qy 4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
Db 304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363
Qy 5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
Db 364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423
Qy 5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
Db 424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483
Qy 5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
Db 484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543
Qy 5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
Db 544 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 603
Qy 5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
Db 604 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 663
Qy 5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
Db 664 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 723
Qy 5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
Db 724 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 783
Qy 5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
Db 784 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 843
Qy 5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554
Db 844 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 903
Qy 5555 AAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTG 5614
Db 904 AAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTG 962
Qy 5615 TTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 5674
Db 963 TTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 1022
Qy 5675 CCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 5734
Db 1023 CCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 1082
Qy 5735 ATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 5794
Db 1083 ATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 1142
Qy 5795 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 5854
Db 1143 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 1202
Qy 5855 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 5914
Db 1203 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 1262
Qy 5915 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 5974
Db 1263 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 1322
Qy 5975 TAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 6034
Db 1323 TAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 1382
Qy 6035 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 6094
Db 1383 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 1442
Qy 6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 6154
Db 1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 1502
Qy 6155 TGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 6214
Db 1503 TGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 1562
Qy 6215 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATA 6274
Db 1563 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATA 1622
Qy 6275 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 6334
Db 1623 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 1682
Qy 6335 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 6394
Db 1683 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 1742
Qy 6395 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 6454
Db 1743 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 1802
Qy 6455 TCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 6514
Db 1803 TCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 1862
Qy 6515 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 6574
Db 1863 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 1922
Qy 6575 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 6634
Db 1923 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 1982
Qy 6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694
Db 1983 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 2042
Qy 6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754
Db 2043 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 2102
Qy 6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814
Db 2103 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 2162
Qy 6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874
Db 2163 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 2222
Qy 6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934
Db 2223 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 2282
Qy 6935 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCC 6994

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       2283 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCC 2342

Qy       6995 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGC 7054
                |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       2343 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGC 2402

Qy       7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATG 7114
                |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATG 2462

Qy       7115 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 7174
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2463 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 2522

Qy       7175 CCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGT 7234
                |||||||||||||||||||||||||||||: ||||||||||||||||||||||||||||||||
Db       2523 CCAATTCAGCAGGAAGCAGTTAGAGCGGTC-TCGGCCAACCTCCCCAACAGCACTTAGGT 2581

Qy       7235 TTTCCTGTTGAGATGGGG 7252
                |||||||||||||||||
Db       2582 TTTCCTGTTGAGATGGGG 2599


RESULT 17
AR699755
LOCUS       AR699755                2599 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 1 from patent US 6919438.
ACCESSION   AR699755
VERSION     AR699755.1  GI:75205544
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2599)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 1 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .2599
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             34.1%;  Score 2551.4;  DB 1;  Length 2599;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2552;  Conservative   39;  Mismatches    5;  Indels    2;  Gaps    2;

Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714
                |||   | | |||||||||||||||||||:||||||:||||||||||||||||:|:||
Db          4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy       4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774
                :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123

Qy       4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
                |||||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183

Qy       4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243

Qy       4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303

Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363

Qy       5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
                |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
                |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483

Qy       5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543

Qy       5195 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 5254
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAAC 603

Qy       5255 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGT 5314
                ||||||||||||||||||||||||||||||||||||||||||||||||||||::||||||||||
Db        604 ATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGT 663

Qy       5315 ATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 5374
                |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        664 ATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTA 723

Qy       5375 CCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 5434
                |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACA 783

Qy       5435 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 5494
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 TCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGC 843

Qy       5495 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 5554
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 TAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGG 903

Qy       5555 AAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTG 5614
                |||||||||||||||||||| |||||||:|||||||||||||||||||||||:||:||
Db        904 AAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTG 962

Qy       5615 TTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 5674
                ||:|||:|||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 TTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCT 1022

Qy       5675 CCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 5734
                |||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGT 1082

Qy       5735 ATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 5794
                |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 ATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACT 1142

Qy       5795 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 5854
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 GCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGA 1202

Qy       5855 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 5914
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 TTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAA 1262

Qy       5915 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 5974
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 CTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAG 1322

Qy       5975 TAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 6034
                ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 TAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCT 1382

Qy       6035 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 6094
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 CAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCC 1442
```

```
Qy       6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 6154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCC 1502

Qy       6155 TGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 6214
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 TGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAG 1562

Qy       6215 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATA 6274
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1563 AAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATA 1622

Qy       6275 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 6334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 CCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCA 1682

Qy       6335 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 6394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 TCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCT 1742

Qy       6395 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 6454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1743 GTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCAT 1802

Qy       6455 TCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 6514
             |||||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db       1803 TCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTA 1862

Qy       6515 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 6574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 AGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTG 1922

Qy       6575 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 6634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 CACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTC 1982

Qy       6635 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 6694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 AAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAAC 2042

Qy       6695 TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTG 6754
             |||||||||||||||||||||||||:||||||||||||||||||:||||||||||||||||
Db       2043 TTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTG 2102

Qy       6755 ARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 6814
             |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2103 AAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCG 2162

Qy       6815 GAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGC 6874
             ||||||||||||||||||||:|||||||||:||||||||||||||||:||||||:||||||
Db       2163 GAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGC 2222

Qy       6875 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCT 6934
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2223 TTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCT 2282

Qy       6935 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCC 6994
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       2283 TAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCC 2342

Qy       6995 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGC 7054
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2343 TTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGC 2402

Qy       7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATG 7114
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATG 2462

Qy       7115 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 7174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2463 TTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCC 2522

Qy       7175 CCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGT 7234
             ||||||||||||||||||||||||||||||||||||||||||||: |||||||||||||||||||||||||
Db       2523 CCAATTCAGCAGGAAGCAGTTAGAGCGGTC-TCGGCCAACCTCCCCAACAGCACTTAGGT 2581

Qy       7235 TTTCCTGTTGAGATGGGG 7252
             |||||||||||||||||||
Db       2582 TTTCCTGTTGAGATGGGG 2599
```

```
RESULT 18
AX007978
LOCUS       AX007978               2599 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9967395.
ACCESSION   AX007978
VERSION     AX007978.1  GI:9995675
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 1 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .2599
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             34.1%;  Score 2551.4;  DB 1;  Length 2599;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2552;  Conservative   39;  Mismatches    5;  Indels    2;  Gaps    2;
```

```
Qy       4655 CCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAA 4714
             |||   | | ||||||||||||||||||:|||||||:|||||||||||||||||:|:||
Db          4 CCCTGCCTTAATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAA 63

Qy       4715 RACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 4774
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GACTGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCT 123

Qy       4775 GGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 4834
             |||||:|||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        124 GGGTAGATACTTTCACGGGTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAG 183

Qy       4835 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 4894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTAC 243

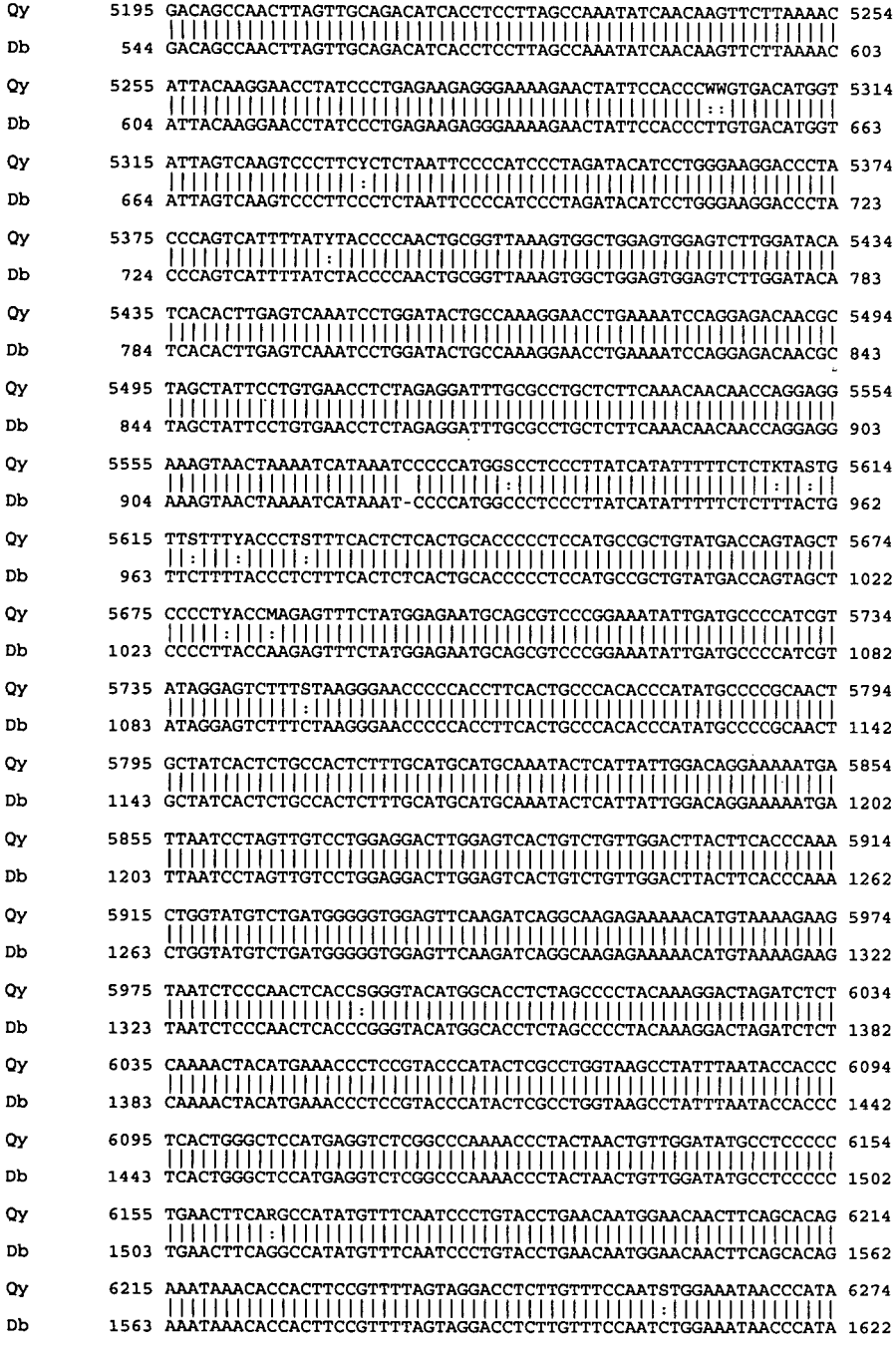
Qy       4895 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 4954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 AGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAG 303

Qy       4955 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 5014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAA 363

Qy       5015 TGAATGAAAYACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 5074
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGAATGAAACACTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGC 423

Qy       5075 CTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 5134
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTA 483

Qy       5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 5194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAA 543
```




RESULT 19
BD221809
LOCUS BD221809 2784 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleic sequence and deduced protein sequence family with human

```
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221809
VERSION     BD221809.1  GI:33031579
KEYWORDS    JP 2002518051-A/4.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2784)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 4 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002518051-A/4
            PD   25-JUN-2002
            PF   23-JUN-1999 JP 2000556036
            PR   23-JUN-1998 FR     98/07920
            PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
            C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
            A61P25/00,
            PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
            sequence and deduced protein sequence family with  CC
            human endogenous
            CC   retroviral motifs, and their uses
            FH   Key             Location/Qualifiers
            FT   source          1. .2784
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2784
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match           30.9%;  Score 2315.4;  DB 1;  Length 2784;
  Best Local Similarity   89.9%;  Pred. No. 0;
  Matches 2523;  Conservative   39;  Mismatches  180;  Indels   66;  Gaps    9;

Qy       4674 CTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTA 4732
             ||||||||||:||||||:||||||||| |||||  :|:||:|||||||||| |||||||
Db          1 CTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAAGAGAAGACTGGCAACTAGATTTTA 60

Qy       4733 CCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGG 4792
             |||| | ||||||| |||||||||||||||||||||||| ||||||:|||||||||| |
Db         61 CCCATATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGATACTTTCACTG 120

Qy       4793 GTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAG 4852
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||      |
Db        121 GTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAA------CG 174

Qy       4853 TTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTG 4912
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db        175 TTCATGAAATAATTCCCAGATTCGGACTTCCCCAAGGCTTACAGAGTGACAATGGCCCTG 234

Qy       4913 CTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTAC 4972
             ||||| |||| |||||||||| |||||||||||| |||||||||| |||||||| ||
Db        235 CTTTCAAGGCTACAGTAACCCAAGGAGTATCCCAGGTGTTAGGTATACAATATCACTCAC 294

Qy       4973 ACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAG 5032
             |||||||||| |||||||||||||||| ||||| |||||||||| ||||:|||||||
Db        295 ACTGCGCCTGGAGGCCACAGTCCTCAGGAAAGGTGGAGAAAATGAACAAAACACTCAAAT 354

Qy       5033 GACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAG 5092
             |||||||||||||| || |||||||||||||||| |||||||||:|||||||||||||||
Db        355 GACATCTAAAAAAGCTAATCCAGGAAACCCACCTCGCATGGCCTGCTCTGTTGCCTATAG 414

Qy       5093 CCTTAAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152
             |||||  ||||||| | |||| |||||||||||||||||||| |||||| |||||||||
Db        415 CCTTACTAAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGTCCATACAAAATGCTGT 474

Qy       5153 ATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTG 5211
             ||||| |||||||| |||||||||| |||| |||||||| | ||||||||||||||||||
Db        475 ATGGACGGCCCTTCCTAACCAATGAACTTGGGCTTGACCGAGAGACAGCCAACTTAGTTG 534

Qy       5212 CAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATC 5271
             |||||||| |||||||||||||||||||| |||||||||||||||| ||| ||
Db        535 CAGACATCATCTCCTTAGCCAAATATCAACAGGTTCTTAAAACATTACAGGGAGCCTGTC 594

Qy       5272 CCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTC 5331
             ||  |||||||||||| |||||||||||||: |||||||||||||||||||||||||||
Db        595 CCCAAGAAGAGGGAAAGGAACTATTCCACCCTGGTGACATGGTATTAGTCAAGTCCCTTC 654

Qy       5332 YCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTA 5391
             :|||||||||||||||||||||||||||||||||||||  |||||||| ||||||||:||
Db        655 CCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGAAACTACCCAGCCATTTTATCTA 714

Qy       5392 CCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAT 5451
             ||| ||| || |||||||||||||||| ||||||||||||||||||||||  |||||||
Db        715 CCCTAACGGCAGTTAAAGTGGCTGGAGCGGAGTCTTGGATACATCACACTCAAGTCAAAC 774

Qy       5452 CCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAAC 5511
             |||||||||||||||||||    |||||||| ||||||| ||||||||||||||||||||
Db        775 CCTGGATACTGCCAAAGGAACTCAAAAATCCATGAGACAATGCTAGCTATTCCTGTGAAC 834

Qy       5512 CTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCAT 5571
             |||||||||| |||||||||||||||  |||||||| |||||||||||||||| |
Db        835 CTCTAGAGGATCTGCGCCTGCTCTTCAAATGACAACCAGGGGGAAAGTAACTAAAATCGT 894

Qy       5572 AAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTC 5631
             ||||||||  |||:|||||||||||||||||||||||:||:||||:| |:|||| :||||
Db        895 AAATCCCC--TGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTCTTACCCCCTTTC 952

Qy       5632 ACTCTCACTGCACCCCCTCCATGCCGCTGTA------------------TGACCAGTAG 5672
             |||||||||||||||| |||||||| ||| |                    |||||||
Db        953 ACTCTCACTGCACCCCGTCCATGCCACTGCACCCCGTCCATGCCCCGTCCATGCCAGTAG 1012

Qy       5673 CTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATC 5732
             |||||||:| |:|||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTCCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATT 1072

Qy       5733 GTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAA 5792
             |||||||||| | |:|||||||||||||||||||||||||||||||||||||||||||| |||
Db       1073 GTATAGGAGTTTATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAA 1132

Qy       5793 CTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAT 5852
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAC 1192

Qy       5853 GATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCA 5912
             |||||||| |||||||||||||||||||||              |||||| ||||||| ||
Db       1193 GATTAATCCCAGTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCA 1239

Qy       5913 AACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGA 5972
               ||  |||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db       1240 TACCAGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGA 1299

Qy       5973 AGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCT 6032
             |||||||||||||||| ||| |||||||| ||||| || ||||||||||||||||||||||
Db       1300 AGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCT 1359

Qy       6033 CTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCAC 6092
             |||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db       1360 CTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCAC 1419

Qy       6093 CCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCC 6152
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 CCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCC 1479

Qy       6153 CCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCAC 6212
             |||| |||| |:|||||  |||||||||| ||||||||||||||||||||||||||||||
Db       1480 CCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCAC 1539

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             |||||||||||||||| |||||||||| |||| |||||||:||||||||||||||
```

```
Db       1540 AGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCA 1596

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332

Db       1597 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATG 1656

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392

Db       1657 CATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 1716

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452

Db       1717 CTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTC 1776

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512

Db       1777 ATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACC 1836

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572

Db       1837 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 1896

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632

Db       1897 CGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTC 1956

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692

Db       1957 TCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCA 2016

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGC 6752

Db       2017 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGC 2076

Qy       6753 TGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATC 6812

Db       2077 GGAAAGCGGGGGAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATC 2136

Qy       6813 CGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGA 6872

Db       2137 CGGAATCATCACCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGA 2196

Qy       6873 GCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTT 6932

Db       2197 GCTGCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTT 2256

Qy       6933 CTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCT 6992

Db       2257 CTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCT 2316

Qy       6993 CCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAAT 7040

Db       2317 CCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAAT 2376

Qy       7041 GGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAG 7100

Db       2377 GGAGCCCCAGATGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAG 2436

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160

Db       2437 CCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACA 2496

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220

Db       2497 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC 2556

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280

Db       2557 AACAGCAGTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTC 2616

Qy       7281 CTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACA 7340

Db       2617 CTAGACCAACTAAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACA 2676
```

```
Qy       7341 CGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAAT 7391

Db       2677 CCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAAT 2736

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTG 7439

Db       2737 GCTAATTAGGCAAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTG 2784


RESULT 20
AR699758
LOCUS       AR699758               2784 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 4 from patent US 6919438.
ACCESSION   AR699758
VERSION     AR699758.1  GI:75205547
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2784)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 4 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .2784
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match           30.9%;  Score 2315.4;  DB 1;  Length 2784;
  Best Local Similarity   89.9%;  Pred. No. 0;
  Matches 2523;  Conservative   39;  Mismatches  180;  Indels   66;  Gaps    9;

Qy       4674 CTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTA 4732

Db          1 CTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAAGAGAAGACTGGCAACTAGATTTTA 60

Qy       4733 CCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGG 4792

Db         61 CCCATATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGATACTTTCACTG 120

Qy       4793 GTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAG 4852

Db        121 GTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAA------CG 174

Qy       4853 TTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTG 4912

Db        175 TTCATGAAATAATTCCCAGATTCGGACTTCCCCAAGGCTTACAGAGTGACAATGGCCCTG 234

Qy       4913 CTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTAC 4972

Db        235 CTTTCAAGGCTACAGTAACCCAAGGAGTATCCCAGGTGTTAGGTATACAATATCACTCAC 294

Qy       4973 ACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAG 5032

Db        295 ACTGCGCCTGGAGGCCACAGTCCTCAGGAAAGGTGGAGAAAATGAACAAAACACTCAAAT 354

Qy       5033 GACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAG 5092

Db        355 GACATCTAAAAAAGCTAATCCAGGAAACCCACCTCGCATGGCCTGCTCTGTTGCCTATAG 414

Qy       5093 CCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152

Db        415 CCTTACTAAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGTCCATACAAAATGCTGT 474

Qy       5153 ATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTG 5211

Db        475 ATGGACGGCCCTTCCTAACCAATGAACTTGGGCTTGACCGAGAGACAGCCAACTTAGTTG 534
```



```
Db       2677 CCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAAT 2736

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTG 7439.
             ||||||||||||||| ||||||| ||||||||||||||||||:| |||
Db       2737 GCTAATTAGGCAAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTG 2784


RESULT 21
AX007981
LOCUS       AX007981               2784 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 4 from Patent WO9967395.
ACCESSION   AX007981
VERSION     AX007981.1  GI:9995678
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 4 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .2784
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             30.9%;  Score 2315.4;  DB 1;  Length 2784;
  Best Local Similarity   89.9%;  Pred. No. 0;
  Matches 2523;  Conservative   39;  Mismatches  180;  Indels   66;  Gaps    9;

Qy       4674 CTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATTTTA 4732
             ||||||||||:|||||:|||||||| |||||  :|:||:|||||||||| |||||||
Db          1 CTCCTTCAGGAGAACAAAGAACAGGCCACTACCCAAGAGAAGACTGGCAACTAGATTTTA 60

Qy       4733 CCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGG 4792
             |||| | ||||||| |||||||||||||||||||||||| ||||||:|||||||||| |
Db         61 CCCATATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTTTGGGTAGATACTTTCACTG 120

Qy       4793 GTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAG 4852
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||      |
Db        121 GTTGGGCAGAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAA------CG 174

Qy       4853 TTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTG 4912
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||
Db        175 TTCATGAAATAATTCCCAGATTCGGACTTCCCCAAGGCTTACAGAGTGACAATGGCCCTG 234

Qy       4913 CTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTAC 4972
             ||||| |||| |||||||||| |||||||||||| |||||||||| |||||||| ||
Db        235 CTTTCAAGGCTACAGTAACCCAAGGAGTATCCCAGGTGTTAGGTATACAATATCACTCAC 294

Qy       4973 ACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAG 5032
             |||||||||| |||||||||||||||| ||||| |||||||||||  |||:|||||||
Db        295 ACTGCGCCTGGAGGCCACAGTCCTCAGGAAAGGTGGAGAAAATGAACAAAACACTCAAAT 354

Qy       5033 GACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAG 5092
             |||||||||||||| || |||||||||||||||| |||||||||:|||||||||||||||
Db        355 GACATCTAAAAAAGCTAATCCAGGAAACCCACCTCGCATGGCCTGCTCTGTTGCCTATAG 414

Qy       5093 CCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152
             |||||  ||||||| | |||| |||||||||||||||||||| |||||| |||||||||
Db        415 CCTTACTAAGAATCCGAAACTCTCCCCAAAAAGCAGGACTTAGTCCATACAAAATGCTGT 474

Qy       5153 ATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAGTTG 5211
             ||||| |||||||| |||||||||| |||| |||||||| | ||||||||||||||||||
Db        475 ATGGACGGCCCTTCCTAACCAATGAACTTGGGCTTGACCGAGAGACAGCCAACTTAGTTG 534

Qy       5212 CAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATC 5271
             |||||||| |||||||||||||||||||||||||||||||||||||||||| ||| ||| ||
Db        535 CAGACATCATCTCCTTAGCCAAATATCAACAGGTTCTTAAAACATTACAGGGAGCCTGTC 594

Qy       5272 CCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTC 5331
             ||  |||||||||||| ||||||||||||||: ||||||||||||||||||||||||||||
Db        595 CCCAAGAAGAGGGAAAGGAACTATTCCACCCTGGTGACATGGTATTAGTCAAGTCCCTTC 654

Qy       5332 YCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTA 5391
             :|||||||||||||||||||||||||||||||||||||||  |||||||| ||||||||:||
Db        655 CCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGAAACTACCCAGCCATTTTATCTA 714

Qy       5392 CCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAT 5451
             ||| ||| || |||||||||||||||| ||||||||||||||||||||  |||||||
Db        715 CCCTAACGGCAGTTAAAGTGGCTGGAGCGGAGTCTTGGATACATCACACTCAAGTCAAAC 774

Qy       5452 CCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAAC 5511
             |||||||||||||||||||||    |||||||| ||||||| ||||||||||||||||||||
Db        775 CCTGGATACTGCCAAAGGAACTCAAAAATCCATGAGACAATGCTAGCTATTCCTGTGAAC 834

Qy       5512 CTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCAT 5571
             |||||||||| |||||||||||||||||  |||||||| |||||||||||||||||| |
Db        835 CTCTAGAGGATCTGCGCCTGCTCTTCAAATGACAACCAGGGGGAAAGTAACTAAAATCGT 894

Qy       5572 AAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTC 5631
             ||||||||  |||:|||||||||||||||||||||:||:||||:| |:|||| :||||
Db        895 AAATCCCC--TGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTCTTACCCCCTTTC 952

Qy       5632 ACTCTCACTGCACCCCCTCCATGCCGCTGTA------------------TGACCAGTAG 5672
             |||||||||||||||||| ||||||| ||| |                  |||||||
Db        953 ACTCTCACTGCACCCCGTCCATGCCACTGCACCCCGTCCATGCCCCGTCCATGCCAGTAG 1012

Qy       5673 CTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATC 5732
             |||||||:| |:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CTCCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATT 1072

Qy       5733 GTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAA 5792
             |||||||||| | |:|||||||||||||||||||||||||||||||||||||||||| |||
Db       1073 GTATAGGAGTTTATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAA 1132

Qy       5793 CTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAT 5852
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAAC 1192

Qy       5853 GATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCA 5912
             |||||||| |||||||||||||||||||||||||||||  ||||| ||||||| ||
Db       1193 GATTAATCCCAGTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCA 1239

Qy       5913 AACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGA 5972
              ||  |||||||||||||||||||||||||||||||||||||||||||||||||||| |||| ||
Db       1240 TACCAGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGA 1299

Qy       5973 AGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCT 6032
             |||||||||||||||| ||| |||||||| ||||| || ||||||||||||||||||||||
Db       1300 AGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCT 1359

Qy       6033 CTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCAC 6092
             |||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db       1360 CTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCAC 1419

Qy       6093 CCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCC 6152
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1420 CCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCC 1479

Qy       6153 CCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCAC 6212
             |||| |||| |:|||||  |||||||||| |||||||||||||||||||||||||||||||
Db       1480 CCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCAC 1539

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             |||||||||||||||||| |||||||||| ||||   ||||||||:|||||||||||||||
Db       1540 AGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCA 1596

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             |||||||||||||||||||||||||||||||||||||||||| | ||||| ||||||||||||||
```

```
Db       1597 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATG 1656

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
              ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1657 CATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 1716

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
              |||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db       1717 CTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTC 1776

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
              |||||||||| ||||:||| |||||||||||||||||||||||||| | |||||||| |
Db       1777 ATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACC 1836

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1837 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 1896

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
              | | ||| ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1897 CGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTC 1956

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
              |||||||||||||| ||||||||| ||||||| ||  ||||||||||||||||||||||
Db       1957 TCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCA 2016

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGC 6752
              ||||||||||||||||||||||||||:||||||||||||||||||:|||||||||||||
Db       2017 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGC 2076

Qy       6753 TGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATC 6812
              ||:|| |||||||||| |||||||||| |||| |||||||||| ||||||||||||||||
Db       2077 GGAAAGCGGGGGAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATC 2136

Qy       6813 CGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGA 6872
              ||||||| |||| |||||||||:||||||||| || ||||||| ||||||:|||| |:||
Db       2137 CGGAATCATCACCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGA 2196

Qy       6873 GCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTT 6932
              ||| | |||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       2197 GCTGCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTT 2256

Qy       6933 CTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCT 6992
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       2257 CTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCT 2316

Qy       6993 CCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAAT 7040
              |||||||| ||||||| |||||||||||||| ||:|||||             |||||
Db       2317 CCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAAT 2376

Qy       7041 GGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAG 7100
              ||||||| |||||||||||| || ||| ||||||| | ||||||||||||||||||:|||
Db       2377 GGAGCCCCAGATGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAG 2436

Qy       7101 CCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACA 7160
              |||| | |||||||||||||||||||||||||||||||||||| |||||||||||| |||||||
Db       2437 CCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACA 2496

Qy       7161 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCC 7220
              |||||||||||||||||||||||||||||||||||||||||| ||| || ||||||||||||||
Db       2497 ACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC 2556

Qy       7221 AACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTC 7280
              ||||||| || |||||||||||||| |||||||||||||||||| || || ||||||
Db       2557 AACAGCAGTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTC 2616

Qy       7281 CTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACA 7340
              |||| |  |:|||||||||:|||| |||||:|||||||||||| | ||||||||||||
Db       2617 CTAGACCAACTAAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACA 2676

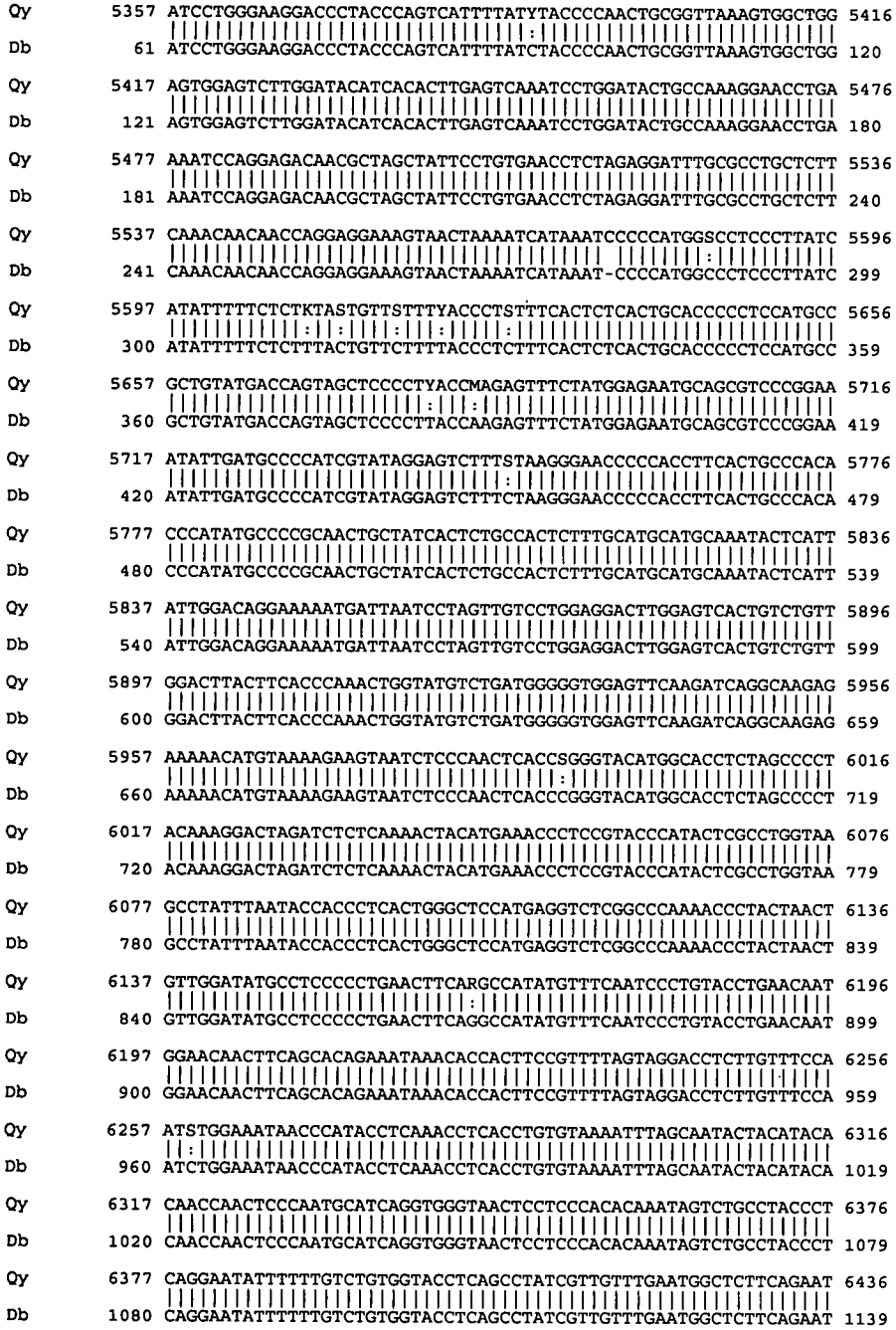
Qy       7341 CGGGGCTTGCAACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAAT 7391
              | |||||||||||||||:|||| ||  ||||||         ||||||||||||||||||
Db       2677 CCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAAT 2736

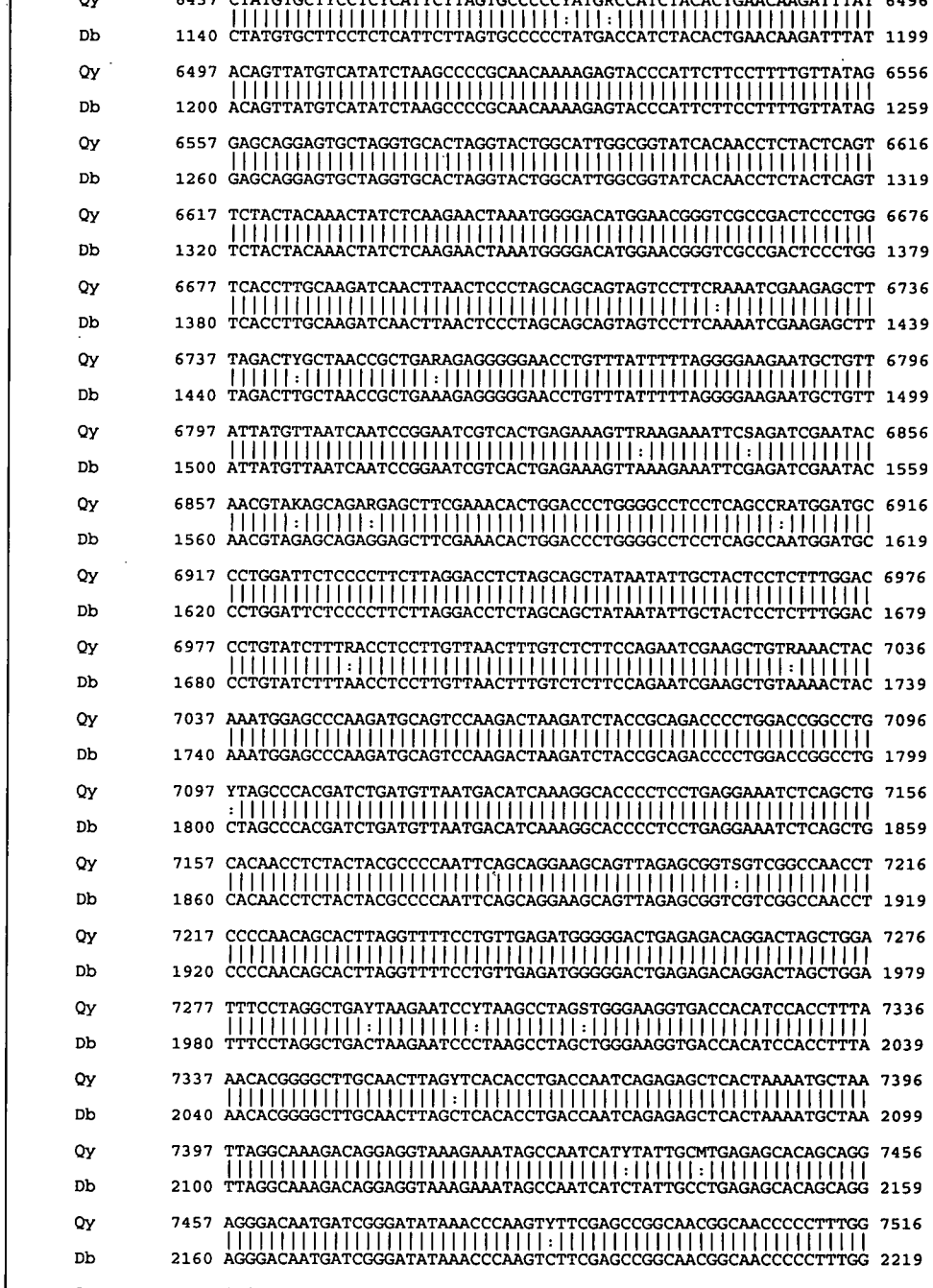
Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTG 7439
              ||||||||||||||| ||||||| ||||||||||||||||||||:| |||
Db       2737 GCTAATTAGGCAAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTG 2784


RESULT 22
AF520483S2
LOCUS       AF520483S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 147 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520484
VERSION     AF520484.1  GI:33410936
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 147 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17564.1"
                     /db_xref="GI:33410937"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60
```





```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
                :||||:
Db       2280 ACTGCA 2285


RESULT 23
AF520485S2
LOCUS       AF520485S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 148 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520486
VERSION     AF520486.1  GI:33410940
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 148 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17565.1"
                     /db_xref="GI:33410941"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;
```

```
Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
              |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
              ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
              ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
```

```
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496

Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556

Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616

Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676

Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736

Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285
```

```
RESULT 24
AF520493S2
LOCUS       AF520493S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 24 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520494
VERSION     AF520494.1  GI:33410956
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 24 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Asian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17569.1"
                     /db_xref="GI:33410957"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
```

```
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db            61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||||:||||||||||||
Db           241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||
Db           300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db           360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db           420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db           660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db           840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||:||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159
```

```
Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 25
AF520495S2
LOCUS       AF520495S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 24 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520496
VERSION     AF520496.1  GI:33410960
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 24 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Asian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17570.1"
                     /db_xref="GI:33410961"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"

     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
              |||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
              ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
              ||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCCACCTTCACTGCCCACA 5776
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019
```

```
Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||:|||||||:|||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 26
AF520497S2
LOCUS       AF520497S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 30 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520498
VERSION     AF520498.1  GI:33410964
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 30 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17571.1"
                     /db_xref="GI:33410965"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQARAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
```

```
                RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
  misc_feature  1705. .1720
                /note="splice acceptor site"
  LTR           1959. .2694

Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
Best Local Similarity 98.3%;  Pred. No. 0;
Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             |||||||::|||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||||:||||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db        1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db        1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||:|||||||
Db        1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||
Db        1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
```

```
                |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
                |||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
                |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
                :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 27
AF520499S2
LOCUS       AF520499S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 30 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520500
VERSION     AF520500.1  GI:33410968
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 30 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17572.1"
                     /db_xref="GI:33410969"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
 Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
 Best Local Similarity   98.3%;  Pred. No. 0;
 Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
                |||||::|||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
                |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
                |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
                ||||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
                |||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
                ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
                ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
                ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
```

```
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

```
RESULT 28
AF520501S2
LOCUS       AF520501S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 34 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520502
VERSION     AF520502.1  GI:33410972
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 34 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17573.1"
                     /db_xref="GI:33410973"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQARERKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNPSTEINTTSV
```

```
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||||:||:|||||:|||:||||||:||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db        1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||||||
Db        1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979
```

```
Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 29
AF520503S2
LOCUS       AF520503S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 34 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520504
VERSION     AF520504.1  GI:33410976
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 34 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17574.1"
                     /db_xref="GI:33410977"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIPFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
```

```
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256

Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316

Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376

Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436

Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496

Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556

Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616

Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676

Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736

Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285
```

```
RESULT 30
AF520507S2
LOCUS       AF520507S2             2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 37 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520508
VERSION     AF520508.1  GI:33410984
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 37 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17576.1"
                     /db_xref="GI:33410985"
```

```
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db        1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:||||||||||||||||||||||||||||||||||||||||:||||||
Db        1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919
```

```
Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 31
AF520513S2
LOCUS       AF520513S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 49 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520514
VERSION     AF520514.1  GI:33410996
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 49 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Asian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17579.1"
                     /db_xref="GI:33410997"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779
```

```
Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:||||||||||:||||||||||:||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 32
AF520515S2
LOCUS       AF520515S2             2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 49 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520516
VERSION     AF520516.1  GI:33411000
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 49 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Asian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
```

```
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17580.1"
                     /db_xref="GI:33411001"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||:||||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db            61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||:||||||||||||
Db           241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db           300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db           360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db           420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db           660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db           840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db          1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db          1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db          1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db          1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||||
Db          1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||:|||||||
Db          1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||:|||||:|||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285


RESULT 33
AF520525S2
LOCUS       AF520525S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 65 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520526
VERSION     AF520526.1  GI:33411020
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 65 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /country="Brazil"
                     /note="Metis
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17585.1"
                     /db_xref="GI:33411021"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||:|||||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
```

```
                ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
                ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
                ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
                ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
                ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
                ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
                ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
                ||||||:||||||:||||||||||||||||||||||||||||||||:||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
                ||||||||||:|||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                ||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
                ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
                ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
                ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
                :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 34
AF520527S2
LOCUS       AF520527S2             2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 65 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520528
VERSION     AF520528.1  GI:33411024
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 65 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
```

```
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /country="Brazil"
                     /note="Metis
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17586.1"
                     /db_xref="GI:33411025"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match            30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity  98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||| |||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:||||||||||||||||||||||||||||||||:||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||||:||||||||||||||||||||||||||||||||||||:|||||||
```

```
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285


RESULT 35
AF520531S2
LOCUS       AF520531S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 78 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520532
VERSION     AF520532.1  GI:33411032
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 78 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17588.1"
                     /db_xref="GI:33411033"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||:|||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 36
AF520533S2
LOCUS       AF520533S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 80 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520534
VERSION     AF520534.1  GI:33411036
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
```

```
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 80 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17589.1"
                     /db_xref="GI:33411037"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||:||||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:|||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619
```

```
Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 37
AF520537S2
LOCUS       AF520537S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 81 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520538
VERSION     AF520538.1  GI:33411044
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 81 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17591.1"
                     /db_xref="GI:33411045"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
```

```
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896

Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956

Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016

Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076

Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136

Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196

Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256

Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316

Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376

Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436

Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496

Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556

Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616

Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676

Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736

Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285
```

RESULT 38
AF520539S2

| | |
|---|---|
| LOCUS | AF520539S2 2694 bp DNA linear PRI 11-FEB-2004 |
| DEFINITION | Homo sapiens individual 81 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence. |
| ACCESSION | AF520540 |
| VERSION | AF520540.1 GI:33411048 |
| KEYWORDS | . |
| SEGMENT | 2 of 2 |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 2694) |
| AUTHORS | Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B. |
| TITLE | The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004) |
| PUBMED | 14757826 |
| REFERENCE | 2 (bases 1 to 2694) |

```
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 81 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17592.1"
                     /db_xref="GI:33411049"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             |||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:|||||:|||:||||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559
```

```
Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

```
RESULT 39
AF520541S2
LOCUS       AF520541S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 82 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520542
VERSION     AF520542.1  GI:33411052
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
    PUBMED  14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 82 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17593.1"
                     /db_xref="GI:33411053"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
```

```
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479
Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539
Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599
Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659
Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719
Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779
Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839
Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899
Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959
Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019
Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079
Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139
Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199
Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259
Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319
Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379
Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439
Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499
Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559
Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619
Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679
Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739
Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799
Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859
Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919
Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979
Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039
Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099
Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159
Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219
Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279
Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

```
RESULT 40
AF520543S2
LOCUS       AF520543S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 82 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520544
VERSION     AF520544.1  GI:33411056
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
```

```
          in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 82 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17594.1"
                     /db_xref="GI:33411057"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||:||||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db         420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db         840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499
```

```
Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 41
AF520547S2
LOCUS       AF520547S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 83 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520548
VERSION     AF520548.1  GI:33411064
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 83 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17596.1"
                     /db_xref="GI:33411065"
                     /translation="MALPYHIPLFTVLLPSFTLTAPPPCRCMTSSSPYQEPLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIPFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCCTCCATGCC 359
```

```
Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             |||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

RESULT 42
AF520549S2

```
LOCUS       AF520549S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 95 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520550
VERSION     AF520550.1  GI:33411068
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 95 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17597.1"
                     /db_xref="GI:33411069"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||
```

```
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439
```

```
Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||:|||||||:|||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285


RESULT 43
AF520551S2
LOCUS       AF520551S2             2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 95 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520552
VERSION     AF520552.1  GI:33411072
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 95 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17598.1"
                     /db_xref="GI:33411073"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.2%;  Score 2258.9;  DB 1;  Length 2694;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2246;  Conservative   39;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:|||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299
```

```
Qy      5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||
Db       300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy      5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db       360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy      5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy      5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy      5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy      5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy      5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy      6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy      6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy      6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy      6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy      6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy      6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy      6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy      6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db      1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy      6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy      6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy      6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy      6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db      1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy      6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db      1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy      6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db      1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy      6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:||||||||||||||||||||||||||||||||||||:|||||||||
Db      1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy      6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy      6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||
Db      1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy      7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy      7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy      7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db      1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy      7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy      7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:||||||||:||||||||:||||||||||||||||||||||||||||||
Db      1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy      7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy      7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||
Db      2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy      7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy      7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy      7577 RCTGCR 7582
             :||||:
Db      2280 ACTGCA 2285
```

RESULT 44
AF520523S2

| | |
|---|---|
| LOCUS | AF520523S2 2694 bp DNA linear PRI 11-FEB-2004 |
| DEFINITION | Homo sapiens individual 6 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence. |
| ACCESSION | AF520524 |
| VERSION | AF520524.1 GI:33411016 |
| KEYWORDS | . |
| SEGMENT | 2 of 2 |

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 6 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17584.1"
                     /db_xref="GI:33411017"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2257.3;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2245;  Conservative   39;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACGCTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||||||| |||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
              ||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
              ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
              ||||||:||||||:||||||||||||||||||||||||||||||||:|||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
              ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              |||||||||||||:|||||||||:||||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
              ||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 45
AF520535S2
LOCUS       AF520535S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 80 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520536
VERSION     AF520536.1  GI:33411040
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 80 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17590.1"
                     /db_xref="GI:33411041"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.1%;  Score 2257.3;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2245;  Conservative   39;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
                                                                          
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
                                                                          
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
                                                                          
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
                                                                          
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
                                                                          
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
                                                                          
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
                                                                          
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
                                                                          
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
                                                                          
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
                                                                          
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
                                                                          
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
                                                                          
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
                                                                          
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
                                                                          
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
                                                                          
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
                                                                          
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
                                                                          
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
                                                                          
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319
```

```
Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
                                                                          
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
                                                                          
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
                                                                          
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
                                                                          
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
                                                                          
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
                                                                          
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
                                                                          
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                                                                          
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                                                                          
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                                                                          
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                                                                          
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                                                                          
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
                                                                          
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
                                                                          
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
                                                                          
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
                                                                          
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
                  :    :
Db       2280 ACTGCA 2285
```

```
RESULT 46
AF520563S2
LOCUS       AF520563S2              2694 bp    DNA     linear   PRI 11-FEB-2004
```

```
DEFINITION  Homo sapiens individual 71 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520564
VERSION     AF520564.1  GI:33411096
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 71 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="African
                     endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17604.1"
                     /db_xref="GI:33411097"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2257.3;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2245;  Conservative   39;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAGCCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||| ||||||||:||||||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             |||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259
```

```
Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 47
BD196244
LOCUS       BD196244               2372 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196244
VERSION     BD196244.1  GI:33006014
KEYWORDS    JP 2002512530-A/10.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2372)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 10 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/10
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2372
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              ||||||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
```

```
Db      420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479
Qy      595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
Db      480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539
Qy      655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
Db      540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599
Qy      715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
Db      600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659
Qy      775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
Db      660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719
Qy      835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
Db      720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779
Qy      895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
Db      780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839
Qy      955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
Db      840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899
Qy     1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
Db      900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958
Qy     1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
Db      959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018
Qy     1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
Db     1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078
Qy     1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
Db     1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137
Qy     1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
Db     1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196
Qy     1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
Db     1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256
Qy     1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
Db     1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312
Qy     1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
Db     1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371
Qy     1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
Db     1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430
Qy     1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
Db     1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490
Qy     1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
Db     1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550
Qy     1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
Db     1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610
Qy     1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
Db     1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670
Qy     1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
Db     1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730
Qy     1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
Db     1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790
Qy     1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
Db     1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850
Qy     1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
Db     1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910
Qy     2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
Db     1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970
Qy     2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
Db     1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030
Qy     2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
Db     2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090
Qy     2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
Db     2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150
Qy     2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
Db     2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210
Qy     2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
Db     2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270
Qy     2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
Db     2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330
Qy     2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
Db     2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

```
RESULT 48
BD267486
LOCUS       BD267486               2372 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267486
VERSION     BD267486.1  GI:33077254
KEYWORDS    JP 2002534980-A/27.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2372)
```

```
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 27 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/27
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .2372
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||:|||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db         959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:|||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              ||||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||
Db        1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||
Db        1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              ||||||||   ||| | |      ||||| ||||||||||||||||||||||||||||||
Db        1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:|||||||||||||||||
Db        1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730
```

```
Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
             |||||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372


RESULT 49
AX000965
LOCUS       AX000965               2372 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 10 from Patent WO9902696.
ACCESSION   AX000965
VERSION     AX000965.1  GI:7241207
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2372)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 10 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy        1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy        1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
               ||||:|||::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy        1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
               ||:|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy        1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
               ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137
```

```
Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
             ||||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
             ||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
             ||||||||:|||||||||||||||||:|||| ||:|||||:||||||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
             |||||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

RESULT 50
AX027479

```
LOCUS       AX027479               2372 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 29 from Patent FR2788784.
ACCESSION   AX027479
VERSION     AX027479.1  GI:10188443
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 29 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

Query Match 30.1%; Score 2256.8; DB 1; Length 2372;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2332; Conservative 30; Mismatches 10; Indels 10; Gaps 8;

```
Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              ||||||||||||||||||||||||||||||||||||||:||| ||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:|||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              ||||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||::||||||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:|||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              ||||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              ||||||||||||||||||||||||||||||||||||:||||||||||||||:||||||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              ||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:||||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy       1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy       1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy       2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy       2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy       2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy       2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy       2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy       2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
              |||||||||||||||||||||||||||||||||||||||||||
Db       2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

```
RESULT 51
AF072504
LOCUS       AF072504               2372 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
ACCESSION   AF072504
VERSION     AF072504.1  GI:4262288
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2372)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
```

```
REFERENCE   2  (bases 1 to 2372)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.44.4"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     misc_feature    1. .2372
                     /note="corresponds to gag and pol"
     repeat_region   1. .2372
                     /rpt_family="HERV-W"
     5'UTR           <1. .6
                     /note="R"
     5'UTR           7. .460
                     /note="U5"
     primer_bind     464. .481
                     /note="putative primer binding site for tRNA-W"

  Query Match             30.1%;  Score 2256.8;  DB 1;  Length 2372;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 2332;  Conservative   30;  Mismatches   10;  Indels   10;  Gaps    8;

Qy        115 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCAC 60

Qy        175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 234
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAG 120

Qy        235 GGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 294
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 GGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAG 180

Qy        295 GCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 354
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTTGCCATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTA 240

Qy        355 NTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTT 414
               ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        241 ATCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTT 300

Qy        415 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGA 474
              |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||
Db        301 ACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGA 359

Qy        475 GAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACC 534
              ||||||||:||||||||||||||||||||||||||||:||||||||||||||||||||
Db        360 GAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCGTCTTGGAAGTGGTTCACC 419

Qy        535 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGG 594
              |||||||||||||||||||||||||||||||||||||:||||||||||:||||:::||||
Db        420 ACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCAACGACGG 479

Qy        595 ACATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 654
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 ACATCCAAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTG 539

Qy        655 GARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 714
              ||:||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GAGAATTGGGACCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGC 599

Qy        715 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        600 AGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCT 659

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 719

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 779

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||::||||||||||||||||||||||||
Db        780 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 839

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 899

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        900 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 958

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1018

Qy       1135 AAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1194
              ||:|||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       1019 AACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGG 1078

Qy       1195 AGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1254
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCAT 1137

Qy       1255 AACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1314
              ||||||||||||||:|||||||| ||||||||||||||||||||||||||||||||||||
Db       1138 AACTGCAGCCTGAGGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATANGGATG 1196

Qy       1315 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCC 1374
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||
Db       1197 ACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCC 1256

Qy       1375 TCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTT 1434
              ||||||||    ||| | |     ||||| ||||||||||||||||||||||||||||||
Db       1257 TCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTT 1312

Qy       1435 GTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCC 1494
              ||||||||:|||||||||||||||||||:|||| ||:||||:|||||||||||||||||
Db       1313 GTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCC 1371

Qy       1495 ACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1554
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGC 1430

Qy       1555 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 ATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCG 1490

Qy       1615 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGC 1550

Qy       1675 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGAT 1734
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1551 CCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGAT 1610

Qy       1735 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCAT 1670

Qy       1795 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 GACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAAT 1730

Qy       1855 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1914
```

```
Db      1731 GCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAG 1790

Qy      1915 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1974
Db      1791 TAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCA 1850

Qy      1975 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 2034
Db      1851 CTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCA 1910

Qy      2035 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 2094
Db      1911 GGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGT 1970

Qy      2095 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2154
Db      1971 ATGCTTGACCATTGAGGGCCAGGAAGGTTGTCTCCTGGACACTGGTGCGGTCTTCTTAGT 2030

Qy      2155 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2214
Db      2031 CTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNT 2090

Qy      2215 AAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2274
Db      2091 AAGACGGGCAGTCACTAGATACTTTTTCCCAGCCACTAAGTTATGAACTGGGGAGCTTTA 2150

Qy      2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
Db      2151 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2210

Qy      2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2394
Db      2211 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGT 2270

Qy      2395 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2454
Db      2271 TGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGG 2330

Qy      2455 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2496
Db      2331 ACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGG 2372
```

```
RESULT 52
AF520477S2
LOCUS       AF520477S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 132 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520478
VERSION     AF520478.1  GI:33410924
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 132 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17561.1"
                     /db_xref="GI:33410925"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
Best Local Similarity   98.2%;  Pred. No. 0;
Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
Db         1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy      5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
Db        61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy      5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
Db       121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy      5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
Db       181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy      5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
Db       241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy      5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
Db       300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy      5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
Db       360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy      5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
Db       420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy      5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
Db       480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy      5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
```

```
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956

Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016

Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076

Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136

Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196

Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256

Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316

Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376

Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436

Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496

Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556

Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616

Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676

Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736

Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285
```

```
RESULT 53
AF520479S2
LOCUS       AF520479S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 132 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520480
VERSION     AF520480.1  GI:33410928
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
```

```
          d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 132 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17562.1"
                     /db_xref="GI:33410929"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTQVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVFSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match          30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity  98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||| ||||||||:||||||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCAGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:|||||||||||||||||||||||||||||||||||:|||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619
```

```
Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||| ||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTTTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:|||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 54
AF520487S2
LOCUS       AF520487S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 148 allele B, envelope glycoprotein gene,
            complete cds, and 3′ long terminal repeat, complete sequence.
ACCESSION   AF520488
VERSION     AF520488.1  GI:33410944
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 148 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17566.1"
                     /db_xref="GI:33410945"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694
```

```
Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
Best Local Similarity   98.2%;  Pred. No. 0;
Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479
```

```
Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

```
RESULT 55
AF520505S2
LOCUS       AF520505S2            2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 37 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520506
VERSION     AF520506.1  GI:33410980
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
```

```
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 37 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17575.1"
                     /db_xref="GI:33410981"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             |||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
```

```
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
              ||||||:||||||:||||||||||||||||||||||||||||||||||||:||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
              ||||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              ||||||||||||:||||||||:|||||||||:||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
              |||||||| ||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 56
AF520509S2
LOCUS       AF520509S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 45 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520510
VERSION     AF520510.1  GI:33410988
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 45 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17577.1"
                     /db_xref="GI:33410989"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::||||||||||||||||||||||:||||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAGCCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
              |||||||||||||||||| |||||||||||||||| |||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAGGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
              ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
              ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419
```

```
Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||:||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||:||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:||||||||:||||||||:||||||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 57
AF520511S2
LOCUS       AF520511S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 45 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520512
VERSION     AF520512.1  GI:33410992
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
```

```
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 45 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17578.1"
                     /db_xref="GI:33410993"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAGCCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||| |||||||||||||||| |||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAGGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
```

```
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||:||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||||:||||||:||||||||||||||
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
             :||||:
Db       2280 ACTGCA 2285


RESULT 58
AF520517S2
LOCUS       AF520517S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 55 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520518
VERSION     AF520518.1  GI:33411004
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.

REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 55 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17581.1"
                     /db_xref="GI:33411005"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::||||||||||||||||||||||:||||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359
```

```
Qy      5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db       360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy      5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy      5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy      5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy      5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy      5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy      6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy      6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy      6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy      6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy      6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy      6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy      6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy      6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db      1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy      6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy      6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy      6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy      6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db      1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy      6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy      6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db      1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy      6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:|||||||:|||||||||||||||||||||||||||||||||||:||||||||
Db      1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy      6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy      6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||||
Db      1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy      7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy      7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy      7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db      1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy      7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy      7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||
Db      1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy      7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy      7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             |||||||| ||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db      2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy      7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db      2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy      7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy      7577 RCTGCR 7582
             :||||:
Db      2280 ACTGCA 2285
```

```
RESULT 59
AF520519S2
LOCUS       AF520519S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 55 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520520
VERSION     AF520520.1  GI:33411008
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 55 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17582.1"
                     /db_xref="GI:33411009"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
```

```
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285


RESULT 60
AF520529S2
LOCUS       AF520529S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 78 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520530
VERSION     AF520530.1  GI:33411028
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 78 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17587.1"
                     /db_xref="GI:33411029"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356

Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416

Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476

Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536

Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
```

```
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
              ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
              ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
              ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
              ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
              ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
              ||||||||||:||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              ||||||||||||:||||||||:||||||||:||||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
              |||||||| ||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

RESULT 61
AF520553S2
LOCUS AF520553S2 2694 bp DNA linear PRI 11-FEB-2004
DEFINITION Homo sapiens individual 96 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520554
VERSION AF520554.1 GI:33411076

```
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 96 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17599.1"
                     /db_xref="GI:33411077"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAGCCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| ||||||| |||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAACCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
```

```
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736

Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796

Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856

Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916

Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976

Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036

Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096

Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156

Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216

Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276

Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336

Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396

Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456

Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516

Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576

Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582

Db       2280 ACTGCA 2285


RESULT 62
AF520555S2
LOCUS       AF520555S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 96 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520556
VERSION     AF520556.1  GI:33411080
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 96 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17600.1"
                     /db_xref="GI:33411081"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356

Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416

Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476

Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAGCCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
```

```
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240
Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAACCATAAAT-CCCCATGGCCCTCCCTTATC 299
Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359
Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419
Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479
Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539
Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599
Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659
Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719
Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779
Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839
Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899
Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959
Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019
Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079
Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139
Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199
Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259
Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379
Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439
Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499
Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559
Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619
Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679
Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739
Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799
Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859
Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919
Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979
Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039
Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099
Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159
Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219
Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279
Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

RESULT 63
AF520557S2
LOCUS AF520557S2 2694 bp DNA linear PRI 11-FEB-2004

```
DEFINITION  Homo sapiens individual 31 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520558
VERSION     AF520558.1  GI:33411084
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 31 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Ashkenazi
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17601.1"
                     /db_xref="GI:33411085"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:||||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259
```

```
Qy      6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy      6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy      6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db      1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy      6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db      1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy      6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db      1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy      6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||
Db      1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy      6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy      6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||:|||||||
Db      1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy      7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy      7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy      7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db      1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy      7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy      7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||
Db      1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy      7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy      7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||
Db      2100 CTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy      7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy      7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      2220 GTCCCCTCCCTTTGTGTGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

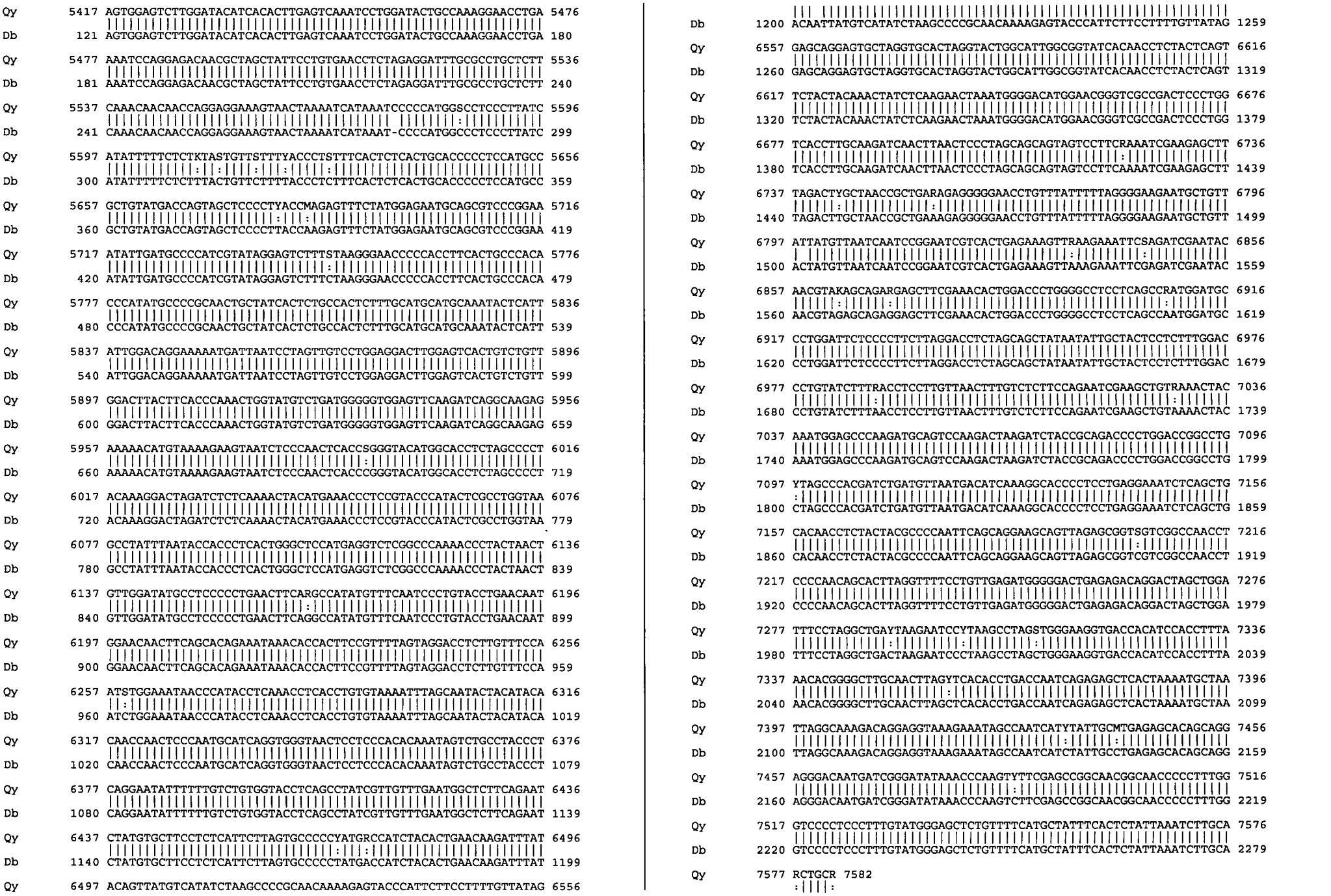
Qy      7577 RCTGCR 7582
             :|||||:
Db      2280 ACTGCA 2285
```

```
RESULT 64
AF520559S2
LOCUS       AF520559S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 31 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520560
VERSION     AF520560.1  GI:33411088
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 31 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Ashkenazi
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17602.1"
                     /db_xref="GI:33411089"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.1%;  Score 2255.7;  DB 1;  Length 2694;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 2244;  Conservative   39;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              ||||||::|||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120
```



```
Db       2280 ACTGCA 2285


RESULT 65
AF520489S2
LOCUS       AF520489S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 22 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520490
VERSION     AF520490.1  GI:33410948
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 22 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17567.1"
                     /db_xref="GI:33410949"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2254.1;  DB 1;  Length 2694;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches 2243;  Conservative   39;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
```

```
                 ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTAGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
```

```
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
              ||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
              ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
              ||||||:||||||:|||||||||||||||||||||||||||||||||:||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
              ||||||||||:|||||||||||||||||||||||||||||||||||:|||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              |||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
              |||||||| ||||||||||||||||||||||||||:||||||:||||||||||||||||
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

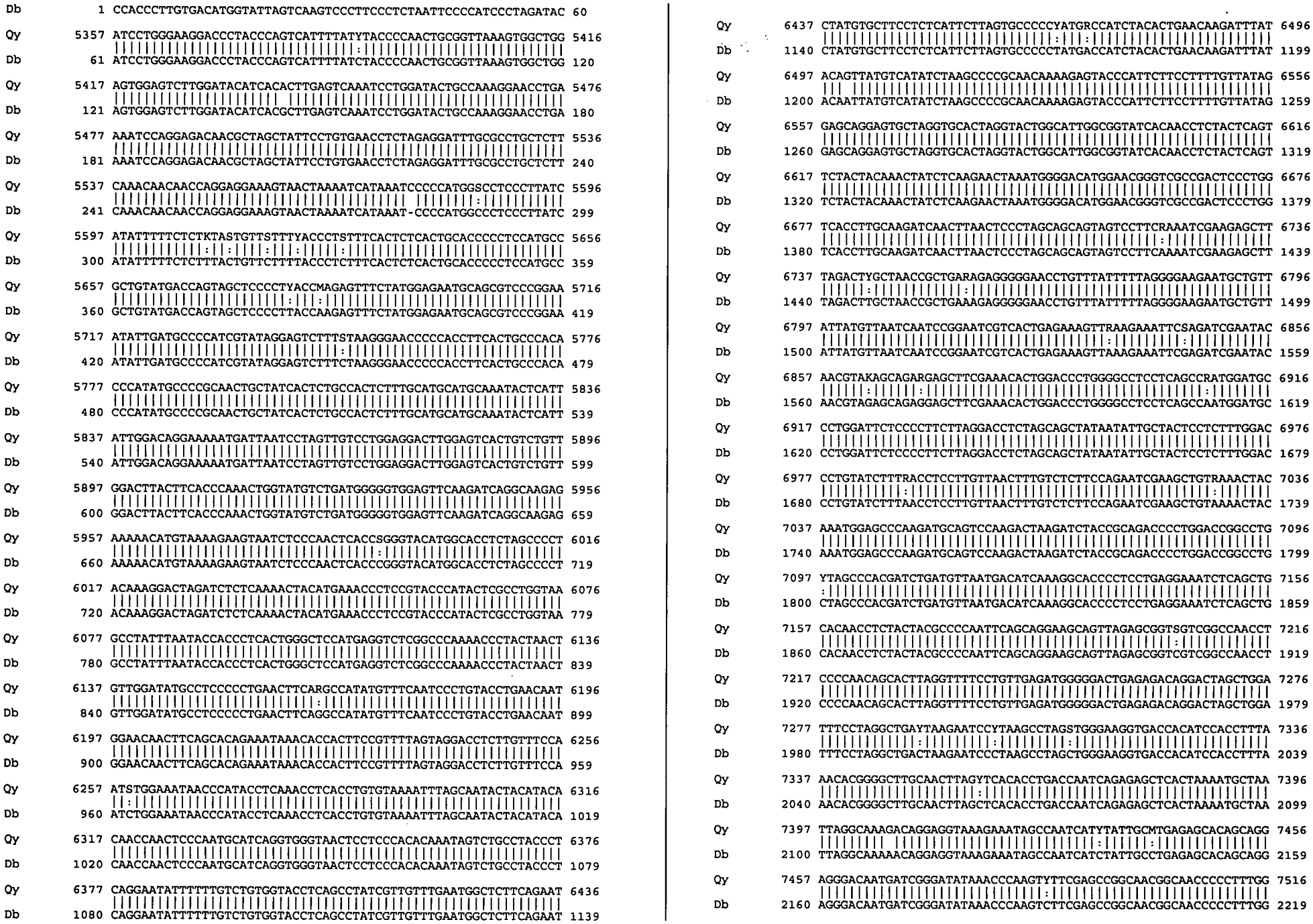
Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

```
RESULT 66
AF520521S2
LOCUS       AF520521S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 6 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520522
VERSION     AF520522.1  GI:33411012
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 6 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17583.1"
                     /db_xref="GI:33411013"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2254.1;  DB 1;  Length 2694;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches 2243;  Conservative   39;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
              |||||::|||||||||||||||||||||||||:|||||||||||||||||||||||||
```



```
Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 67
AF520561S2
LOCUS       AF520561S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 71 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520562
VERSION     AF520562.1  GI:33411092
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 71 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="African
                     endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17603.1"
                     /db_xref="GI:33411093"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match             30.1%;  Score 2254.1;  DB 1;  Length 2694;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches 2243;  Conservative   39;  Mismatches    3;  Indels    1;  Gaps    1;
```

```
Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAGCCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||| |||||||:||||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             |||||||||||||:||:||||:|||:||||||:|||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079
```

```
Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
Db       2280 ACTGCA 2285
```

```
RESULT 68
AF520481S2
LOCUS       AF520481S2             2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 147 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520482
VERSION     AF520482.1  GI:33410932
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 147 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="endogenous virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17563.1"
                     /db_xref="GI:33410933"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
     LTR             1959. .2694

  Query Match           30.1%;  Score 2252.5;  DB 1;  Length 2694;
```

Best Local Similarity 98.1%; Pred. No. 0;
Matches 2242; Conservative 39; Mismatches 4; Indels 1; Gaps 1;

```
Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
Db       1500 ACTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
Db       2100 CTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159
```

```
Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTGTGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 69
AF520545S2
LOCUS       AF520545S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 83 allele A, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520546
VERSION     AF520546.1  GI:33411060
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 83 allele A"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="female"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17595.1"
                     /db_xref="GI:33411061"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    1705. .1720
                     /note="splice acceptor site"
```

```
     LTR             1959. .2694

  Query Match             30.1%;  Score 2252.5;  DB 1;  Length 2694;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches 2242;  Conservative   39;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       5297 CCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATAC 5356
             ||||||::|||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 CCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATAC 60

Qy       5357 ATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGG 5416
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db         61 ATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGG 120

Qy       5417 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 5476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGA 180

Qy       5477 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 5536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTT 240

Qy       5537 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATC 5596
             |||||||||||||||||||||||||||||||||||||||| ||||||||:||||||||||||
Db        241 CAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATC 299

Qy       5597 ATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCC 5656
             ||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db        300 ATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCC 359

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        360 GCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 419

Qy       5717 ATATTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCCACCTTCACTGCCCACA 5776
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        420 ATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCCACCTTCACTGCCCACA 479

Qy       5777 CCCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 5836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATT 539

Qy       5837 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 5896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTT 599

Qy       5897 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 5956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAG 659

Qy       5957 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCT 6016
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        660 AAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCT 719

Qy       6017 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 6076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAA 779

Qy       6077 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 6136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACT 839

Qy       6137 GTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAT 6196
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        840 GTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAAT 899

Qy       6197 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 959

Qy       6257 ATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
             ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 ATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 1019
```

```
Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 1079

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 1139

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       1140 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTAT 1199

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 ACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 1259

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 1319

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 1379

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       1380 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 1439

Qy       6737 TAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 6796
             ||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       1440 TAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTT 1499

Qy       6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATAC 6856
             | ||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
Db       1500 ACTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATAC 1559

Qy       6857 AACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGC 6916
             ||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db       1560 AACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGC 1619

Qy       6917 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 6976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGAC 1679

Qy       6977 CCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTAC 7036
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1680 CCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTAC 1739

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 1799

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 1859

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1860 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCT 1919

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 1979

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             ||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||||||||
Db       1980 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTA 2039

Qy       7337 AACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 7396
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             ||||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||||
Db       2100 CTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTGTGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

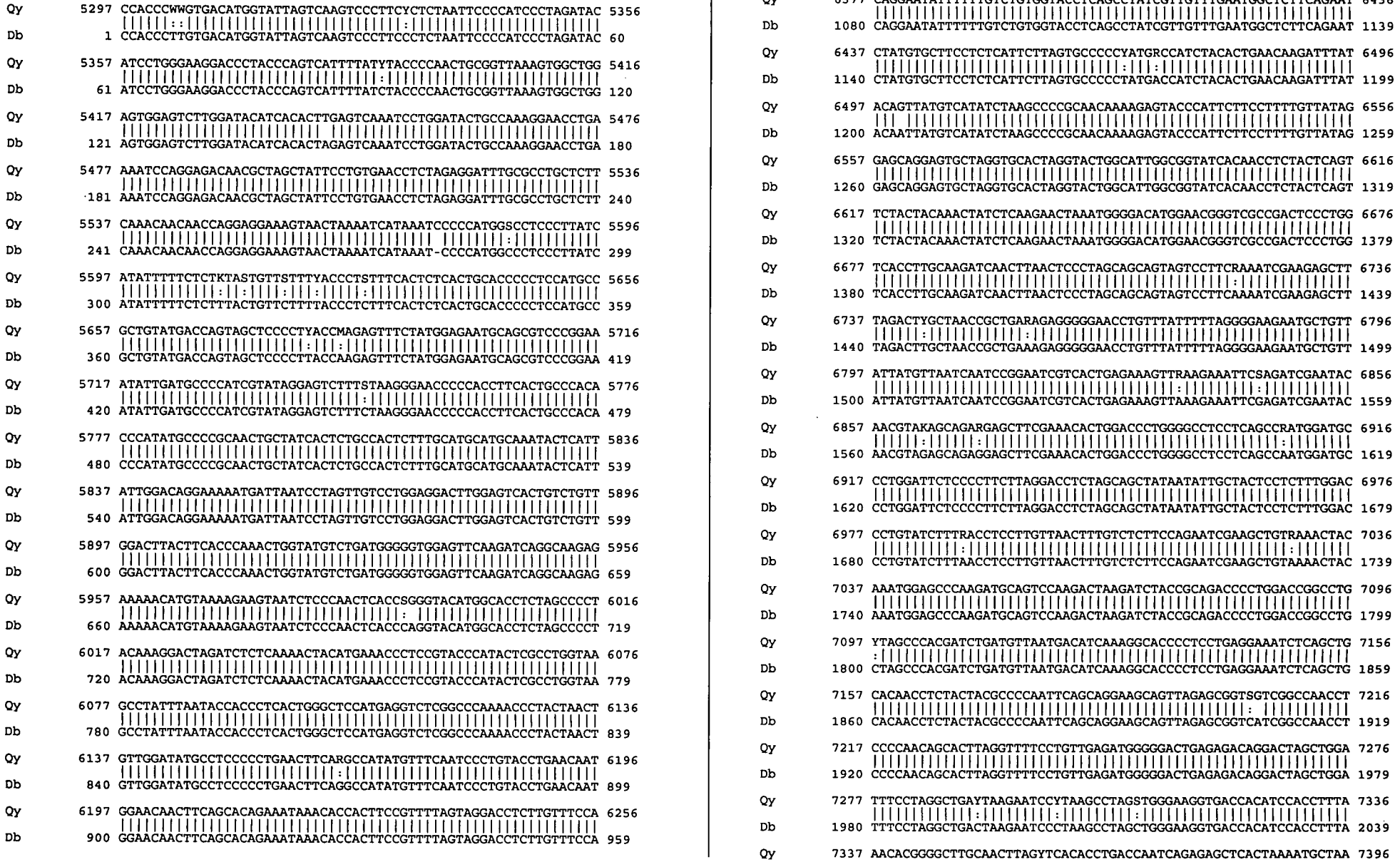
Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285
```

RESULT 70
AF520491S2

```
LOCUS       AF520491S2              2694 bp    DNA     linear   PRI 11-FEB-2004
DEFINITION  Homo sapiens individual 22 allele B, envelope glycoprotein gene,
            complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION   AF520492
VERSION     AF520492.1  GI:33410952
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
  TITLE     The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
   PUBMED   14757826
REFERENCE   2  (bases 1 to 2694)
  AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Superieure de Lyon - 46 allee
            d'Italie, Lyon 69364 cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .2694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="individual 22 allele B"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21-q22"
                     /sex="male"
                     /cell_type="PBMC's"
                     /note="Caucasian
                     endogenous_virus: HERV-W"
     misc_feature    43. .57
                     /note="splice acceptor site"
     CDS             284. .1900
                     /note="syncytin"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17568.1"
                     /db_xref="GI:33410953"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTQVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
```

RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
misc_feature 1705. .1720
/note="splice acceptor site"
LTR 1959. .2694

Query Match 30.1%; Score 2250.9; DB 1; Length 2694;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2241; Conservative 39; Mismatches 5; Indels 1; Gaps 1;



```
Db       2040 AACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAA 2099

Qy       7397 TTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGG 7456
             |||||||| ||||||||||||||||||||||||||||:||||||:|||||||||||||||
Db       2100 TTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGG 2159

Qy       7457 AGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGG 7516
             |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       2160 AGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCCTTTGG 2219

Qy       7517 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 7576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCA 2279

Qy       7577 RCTGCR 7582
              :||||:
Db       2280 ACTGCA 2285


RESULT 71
BD196241
LOCUS       BD196241               2782 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196241
VERSION     BD196241.1  GI:33006011
KEYWORDS    JP 2002512530-A/7.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2782)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 7 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/7
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2782
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:|||||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764
```

RESULT 72
BD267483

```
LOCUS       BD267483               2782 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267483
VERSION     BD267483.1  GI:33077251
KEYWORDS    JP 2002534980-A/24.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2782)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 24 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/24
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .2782
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013
```

```
Qy      5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy      5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy      5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db      1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy      6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy      6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy      6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db      1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy      6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy      6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             .|||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy      6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy      6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy      6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db      1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy      6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy      6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy      6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy      6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db      1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy      6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db      1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy      6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db      1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy      6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||||:|||||
Db      2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy      6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy      6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||:||||
Db      2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy      7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy      7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy      7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db      2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy      7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy      7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:||||||||:||||||||||||||||||||||
Db      2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy      7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db      2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy      7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||
Db      2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy      7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db      2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy      7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy      7572 TTGCARCTGCR 7582
             |||||:||||:
Db      2754 TTGCAACTGCA 2764
```

```
RESULT 73
AX000962
LOCUS       AX000962               2782 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 7 from Patent WO9902696.
ACCESSION   AX000962
VERSION     AX000962.1  GI:7241204
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2782)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 7 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
```

```
 Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1134 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:|||||:|||||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||:|||||||:||||||||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693
```

```
Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764


RESULT 74
AX027476
LOCUS       AX027476               2782 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 26 from Patent FR2788784.
ACCESSION   AX027476
VERSION     AX027476.1  GI:10188440
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 26 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:||||||||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1733

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1793

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1853

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1854 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCGAAATCGAAG 1913

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1914 AGCTTTAGACTTGCTAACCGCTGAGAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1973

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1974 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGAAATTCCAGATCG 2033

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:|||||:||||||||||||||||||||||||||||||||||||:|||
Db       2034 AATACAACGTATAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCGATG 2093

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2153

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||
Db       2154 TGGACCCTGTATCTTTGACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTGAA 2213
```

```
Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2273

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2333

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2334 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCGGCC 2393

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2453

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2454 CTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCTAGGTGGGAAGGTGACCACATCCAC 2513

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       2514 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2573

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||
Db       2574 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTGAGAGCACA 2633

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2634 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCC 2693

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2753

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2754 TTGCAACTGCA 2764


RESULT 75
AX133396
LOCUS       AX133396               2782 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 15 from Patent WO0131021.
ACCESSION   AX133396
VERSION     AX133396.1  GI:14139658
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Conrad,B. and Mach,B.
  TITLE     Multiple sclerosis-related superantigen
  JOURNAL   Patent: WO 0131021-A 15 03-MAY-2001;
            UNIVERSITE DE GENEVE (CH)
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             29.4%;  Score 2205.1;  DB 1;  Length 2782;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2194;  Conservative   36;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        535 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 594

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 654

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 714

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        715 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGGCCTCCC 773

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        774 TTATCATATTTTTCTCTGTAGTGTTCTTTCACCCTGTTTCACTCTCACTGCACCCCCTCC 833

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        834 ATGCCGCTGTATGACCAGTAGCTCCCCTCACCCAGAGTTTCTATGGAGAATGCAGCGTCC 893

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        894 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 953

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1013

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1073

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1133

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1134 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCGGGGTACATGGCACCTCTAG 1193

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1253

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1313

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1314 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1373

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1433

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 TTCCAATGTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1493

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1553

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1613

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db       1614 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1673

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
```

```
                ||||||
```

```
Qy      5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db      1133 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1192

Qy      6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1193 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1252

Qy      6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1253 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1312

Qy      6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db      1313 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1372

Qy      6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1373 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1432

Qy      6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1433 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1492

Qy      6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1493 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1552

Qy      6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1553 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1612

Qy      6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||
Db      1613 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1672

Qy      6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1673 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1732

Qy      6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1733 TATAGGAGCGGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1792

Qy      6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1793 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1852

Qy      6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             |||||||||||||||||||| ||||||||||||||||||||||||||:|||||||||||
Db      1853 CCTGGTCACCTTGCAAGATCAGCTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1912

Qy      6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||
Db      1913 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1972

Qy      6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||||
Db      1973 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2032

Qy      6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||||
Db      2033 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2092

Qy      6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2093 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2152

Qy      6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||||
Db      2153 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2212

Qy      7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2213 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2272

Qy      7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2273 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2332

Qy      7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db      2333 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2392

Qy      7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2393 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2452

Qy      7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db      2453 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2512

Qy      7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db      2513 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2572

Qy      7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||
Db      2573 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2632

Qy      7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db      2633 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2692

Qy      7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2693 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2752

Qy      7572 TTGCARCTGCR 7582
             |||||:||||:
Db      2753 TTGCAACTGCA 2763
```

```
RESULT 77
AF072506
LOCUS       AF072506               2781 bp    mRNA    linear   PRI 11-MAY-1999
DEFINITION  Homo sapiens endogenous retrovirus W envelope protein precursor
            mRNA, complete cds.
ACCESSION   AF072506
VERSION     AF072506.2  GI:4773879
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2781)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 2781)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
REFERENCE   3  (bases 1 to 2781)
  AUTHORS   Bouton,O., Blond,J.-L. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-1999) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
  REMARK    Sequence update by submitter
COMMENT     On May 11, 1999 this sequence version replaced gi:4262291.
```

```
FEATURES             Location/Qualifiers
     source          1. .2781
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.PH74"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     repeat_region   1. .2781
                     /rpt_family="HERV-W"
     5'UTR           <1. .49
                     /note="R"
     5'UTR           50. .503
                     /note="U5"
     primer_bind     507. .524
                     /note="putative primer binding site for tRNA-W"
     misc_feature    534. .535
                     /note="putative splice junction"
     CDS             762. .2378
                     /codon_start=1
                     /product="envelope protein precursor"
                     /protein_id="AAD14546.2"
                     /db_xref="GI:4773880"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     sig_peptide     762. .824
     mat_peptide     825. .2375
                     /product="envelope protein"
     misc_feature    1701. .1712
                     /note="furin cleavage site"
     misc_feature    1713. .2375
                     /note="transmembrane-region site"
     misc_feature    1890. .1949
                     /note="Region: immunosuppressive region"
     3'UTR           2347. .2763
                     /note="U3-R junction undetermined"
     misc_feature    2425. .2435
                     /note="polypurine tract"
     polyA_signal    2745. .2750
     polyA_site      2762

  Query Match             29.4%;  Score 2201.9;  DB 1;  Length 2781;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2192;  Conservative   36;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        534 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 593

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 653

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 713

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        714 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 772

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:|||||:|||:||||||:||||||||||||||||||||||
Db        773 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 832
```

```
Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        833 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 892

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        893 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 952

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 1012

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1072

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1132

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db       1133 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1192

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1252

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1312

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1313 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1372

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1432

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1492

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1552

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1612

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db       1613 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGA 1672

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1732

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 TATAGGAGCGGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1792

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1852

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             |||||||||||||||||||| ||||||||||||||||||||||||||||:|||||||||
Db       1853 CCTGGTCACCTTGCAAGATCAGCTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1912
```

```
Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1913 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1972

Qy       6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCG 6851
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db       1973 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2032

Qy       6852 AATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATG 6911
             ||||||||||:||||||:||||||||||||||||||||||||||||||||||:|||||||
Db       2033 AATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATG 2092

Qy       6912 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 6971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTT 2152

Qy       6972 TGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAA 7031
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||:||||
Db       2153 TGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAA 2212

Qy       7032 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCG 2272

Qy       7092 GCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 7151
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTC 2332

Qy       7152 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       2333 AGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCC 2392

Qy       7212 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 7271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 AACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAG 2452

Qy       7272 CTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCAC 7331
             ||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||||||
Db       2453 CTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAC 2512

Qy       7332 CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAAT 7391
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db       2513 CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAT 2572

Qy       7392 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACA 7451
             ||||||||||||||||||||||||||||||||||||||||||:||||||:||||||||||
Db       2573 GCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACA 2632

Qy       7452 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCC 7511
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       2633 GCAGGAGGGACAATGATCGGGATATAAACCCAAGTCTTCGAGCCGGCAACGGCAACCCCC 2692

Qy       7512 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 7571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATC 2752

Qy       7572 TTGCARCTGCR 7582
             |||||:||||:
Db       2753 TTGCAACTGCA 2763
```

```
RESULT 78
BD196248
LOCUS       BD196248               2575 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196248
VERSION     BD196248.1  GI:33006018
KEYWORDS    JP 2002512530-A/14.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2575)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 14 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/14
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC           diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2575
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2575
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match           28.6%;  Score 2141.4;  DB 1;  Length 2575;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches 2275;  Conservative  134;  Mismatches  159;  Indels   20;  Gaps   11;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTG-AGYCARTYMTCATAC 2969
             |||||||:|||||||||||||||||||||||||||||||||||||| ||:||:|::||||||
Db          1 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAAGCCAATTCTCATAC 60

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
             |||||||||||| ||||| :|||:|||||||||||||||:||: ||::|||||||||||
Db         61 CTGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACC 120

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
             |||||||||||||||||||||||:|||||:||||||||||:|||||||||||||:||||
Db        121 TTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTT 180

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|: | ||:|||||||||||||||||||||||||||| ||:||||:|||||||
Db        181 TCCAAACCAAAGACCCAGCTCTGCTCACAGCAGGTTAAATACTTGGGGCTAAAATTATCC 240

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||||||:|||||||||||||||||:::||| |||||||||||||||||||||||||:|
Db        241 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCAAGCCTATACTGGCTTATCCTCATCCC 300

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||| |||||||||||||||:|::|||||| ||:|||:|||:|||||| |||:|||||||
Db        301 CAAATCCTAAAGCAACTAAGAGAGTTCCTTAGCATAACAGGTTTCTGCTGAATATGGATT 360

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:||||:|||||||||| :|||||:|||| :|||||||||||||||||||||||
Db        361 -CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGC 419

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::|||  ||||||||||||    ||
Db        420 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAG 477

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             ||| |   ||||||||:|||||||||||:||||||||:|||||||||||| :||: ||:
Db        478 AGGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTAT 537

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||||||||    |:||:||||||:|||||||||||||||| :|| :|||||:||||||
Db        538 GTCACAGAAAAAAACAGGAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTT 597

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
```

```
Db        598 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 657
Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        658 TTTATGGGTAGTGGCGGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATACAGGG 717
Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        718 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 777
Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
Db        778 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGT 837
Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
Db        838 GCAGCAACTGCGCAGTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATG---- 893
Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
Db        894 -AAGATAGAACATAACTGCCAACAAGTAATTTCTCAAACCTAGGCCGCTCGAGGGAACCT 952
Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
Db        953 TTTAGAGGTTCCCTTAACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 1011
Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
Db        1012 TAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGCTCAGTGATAATGGAATACTTGAAA 1071
Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
Db        1072 ATAATCCCTTCATTCCAGGAACCAGCGTTCAGCTGGCAGAATTAATAGCCCTCACTCGGG 1131
Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
Db        1132 CATTAGAATTAGGAGAAGGAAAAAGGGTAAATACACATACAGATTCTAAGTATGTTTACT 1191
Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
Db        1192 TAGTCCTCCGTGCCCACGCAGCAATATGGAGAGAAAGGGAATGCTTAACTTCTGAGGGAA 1251
Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229
Db        1252 CACCTATCAAACATCAGGAAGTTATTAGGAGATTATTATTGGCTATACAGAAACCTAAAG 1311
Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289
Db        1312 AGGTGGCAGTCTTACACTGCTGGGGTGGTCAGAAAGAAAAGGAAAGGGAAATAAAAGGGA 1371
Qy       4290 AYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGC 4349
Db        1372 ACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGC 1431
Qy       4350 TTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAG 4409
Db        1432 TTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAATACTTAG 1491
Qy       4410 CAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAG 4466
Db        1492 AAAAAGAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCTCCCCTCAGGATGGCTAG 1551
Qy       4467 CCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCC 4526
Db        1552 CCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCC 1611
Qy       4527 TTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTA 4586
Db        1612 TTCACCAAACCTTTCACTTAGACATTGATAGCACCCATCAGATGGCCAAATCATTATTTA 1671
Qy       4587 CTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARA 4646
```

```
Db        1672 CTGGACCAGGCCTTTTCAAAACTATCAAGCAGCTAGTCAGGGCCTGTGAAGTGTGCCGAA 1731
Qy       4647 RAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACC 4706
Db        1732 GAAATAATCCCATGCCTTATCACCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACC 1791
Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
Db        1792 CAGGAGAAGRVTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 1851
Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
Db        1852 TACTAGTTTGGGTAGATACTTTCACTGGTTGGGCAGAGACCTTCCCCTGTAAGACAGAAA 1911
Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
Db        1912 AGTCCCAAGAGGTAATAAAGGCATTAGTTCATGAAATAATTCCCAGATTCAGACTTCCCT 1971
Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
Db        1972 GAGGCTTACAGAGTGACAATGGCCCTGCTTTCAAGGCTACAGTAACCCA-GGAGTATCCC 2030
Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
Db        2031 AGGTGTTAGGTATACAATATCACTTACACTGCGCCTGGAGG---CAGTCCTCAGGGAAGG 2087
Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCT-AAAAAAGCAAACCCAGGAAACCCAC 5064
Db        2088 CCGAGAAACTGAATGAAACACTCAAACGACATCTAAAAAAAGCTAACCCAGGAAAACCAC 2147
Qy       5065 CTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAA 5124
Db        2148 CTCACATGGCCTGCTCTGTTGCCTATAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAA 2207
Qy       5125 GCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 5184
Db        2208 GCAGGACTTAGCCCATACGAAATGCTATATGGATAGCCCTTCCTAACCAATGACCTTGTG 2267
Qy       5185 CTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAA 5243
Db        2268 CTTGACTGAGAGAGAGCCAACTTAGTTGCAGACATCACCTCCTTATCCAAATATCAACAA 2327
Qy       5244 GTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCW 5303
Db        2328 GTTCTTAAAACATTACAAGGAGCCTGTCCCCGAGAAGAGGGGAAGGAACTATTCCACCCT 2387
Qy       5304 WGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGG 5363
Db        2388 GGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCTCATTGCCTAGATATATCCTGG 2447
Qy       5364 GAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAG 5423
Db        2448 GAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGCAGTAAAAGTGGCTGGAGTGGAG 2507
Qy       5424 TCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCA 5483
Db        2508 TCTTGGATACATCACACTCGAGTCAAACCCTGGATATTACCAAAGGAACCTGAAAATCCA 2567
Qy       5484 GGAGACAA 5491
Db        2568 GGAGACAA 2575
```

```
RESULT 79
AX000969
LOCUS       AX000969               2575 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 14 from Patent WO9902696.
ACCESSION   AX000969
VERSION     AX000969.1  GI:7241211
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
```

```
REFERENCE   1  (bases 1 to 2575)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 14 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .2575
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             28.6%;  Score 2141.4;  DB 1;  Length 2575;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches 2275;  Conservative  134;  Mismatches  159;  Indels   20;  Gaps   11;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTG-AGYCARTYMTCATAC 2969
Db          1 GGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAAGCCAATTCTCATAC 60

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
Db         61 CTGGACACTCTTCTCCTTTGGTATGTGGATGATTTACTTTTAGCTTCCTGTTCAGAAACC 120

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
Db        121 TTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTT 180

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db        181 TCCAAACCAAAGACCCAGCTCTGCTCACAGCAGGTTAAATACTTGGGGCTAAAATTATCC 240

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db        241 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCAAGCCTATACTGGCTTATCCTCATCCC 300

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        301 CAAATCCTAAAGCAACTAAGAGAGTTCCTTAGCATAACAGGTTTCTGCTGAATATGGATT 360

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        361 -CCCAGGTATGGCAAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCAGAAAGC 419

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        420 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGCCC--TAAAG 477

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        478 AGGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTACTTCGTAT 537

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        538 GTCACAGAAAAAACAGGAAATAGCTCTAGGAGTCCTTACACAAGTCTGAGGGATGAGCTT 597

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        598 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 657

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        658 TTTATGGGTAGTGGCGGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATACAGGG 717

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        718 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 777

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
Db        778 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGT 837

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809

Db        838 GCAGCAACTGCGCAGTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATG---- 893

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
Db        894 -AAGATAGAACATAACTGCCAACAAGTAATTTCTCAAACCTAGGCCGCTCGAGGGAACCT 952

Qy       3870 TYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTG 3929
Db        953 TTTAGAGGTTCCCTTAACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTTG 1011

Qy       3930 TAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAA 3989
Db       1012 TAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGCTCAGTGATAATGGAATACTTGAAA 1071

Qy       3990 GTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGG 4049
Db       1072 ATAATCCCTTCATTCCAGGAACCAGCGTTCAGCTGGCAGAATTAATAGCCCTCACTCGGG 1131

Qy       4050 CACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACC 4109
Db       1132 CATTAGAATTAGGAGAAGGAAAAAGGGTAAATACACATACAGATTCTAAGTATGTTTACT 1191

Qy       4110 TAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAA 4169
Db       1192 TAGTCCTCCGTGCCCACGCAGCAATATGGAGAGAAAGGGAATGCTTAACTTCTGAGGGAA 1251

Qy       4170 CACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAG 4229
Db       1252 CACCTATCAAACATCAGGAAGTTATTAGGAGATTATTATTGGCTATACAGAAACCTAAAG 1311

Qy       4230 AGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289
Db       1312 AGGTGGCAGTCTTACACTGCTGGGGTGGTCAGAAAGAAAAGGAAAGGGAAATAAAAGGGA 1371

Qy       4290 AYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGC 4349
Db       1372 ACTGCCAAGCGGATATTGAAGCCAAAAGAGCCGCAAGGCAGGACCCTCCATTAGAAATGC 1431

Qy       4350 TTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAG 4409
Db       1432 TTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCGGGAAACCAAGCCCCAATACTTAG 1491

Qy       4410 CAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAG 4466
Db       1492 AAAAAGAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCTCCCCTCAGGATGGCTAG 1551

Qy       4467 CCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCC 4526
Db       1552 CCACCGAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCC 1611

Qy       4527 TTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTTA 4586
Db       1612 TTCACCAAACCTTTCACTTAGACATTGATAGCACCCATCAGATGGCCAAATCATTATTTA 1671

Qy       4587 CTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARA 4646
Db       1672 CTGGACCAGGCCTTTTCAAAACTATCAAGCAGCTAGTCAGGGCCTGTGAAGTGTGCCGAA 1731

Qy       4647 RAAATAATCCCCTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACC 4706
Db       1732 GAAATAATCCCATGCCTTATCACCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACC 1791

Qy       4707 CTGRARAARACTGGCAACT-GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATC 4765
Db       1792 CAGGAGAAGRVTGGCAACTAGATTTTACCCACATGCCCAAATCTCAGGGATTTCAGTATC 1851

Qy       4766 TACTAGTCTGGGTARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAA 4825
Db       1852 TACTAGTTTGGGTAGATACTTTCACTGGTTGGGCAGAGACCTTCCCCTGTAAGACAGAAA 1911

Qy       4826 AGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCC 4885
```

```
Db       1912 AGTCCCAAGAGGTAATAAAGGCATTAGTTCATGAAATAATTCCCAGATTCAGACTTCCCT 1971

Qy       4886 GAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCC 4945
             |||||||||||||||||||| |||||||||| |||| |||||||||| ||||||||||
Db       1972 GAGGCTTACAGAGTGACAATGGCCCTGCTTTCAAGGCTACAGTAACCCA-GGAGTATCCC 2030

Qy       4946 AGGCGTTAGGTATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGG 5005
             ||| |||||||||| |||||||||||||||||||| |||   |||||||||||||||||
Db       2031 AGGTGTTAGGTATACAATATCACTTACACTGCGCCTGGAGG---CAGTCCTCAGGGAAGG 2087

Qy       5006 TCGAGAAAATGAATGAAAYACTCAAAGGACATCT-AAAAAAGCAAACCCAGGAAACCCAC 5064
              ||||||| ||||||||:||||||| ||||||| |||||||| |||||||||| ||||
Db       2088 CCGAGAAACTGAATGAAACACTCAAACGACATCTAAAAAAAGCTAACCCAGGAAAACCAC 2147

Qy       5065 CTCACATGGCCTGYTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAA 5124
              |||||||||||||:|||||||||||||||||||  |||||||   |||| ||||||||||
Db       2148 CTCACATGGCCTGCTCTGTTGCCTATAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAA 2207

Qy       5125 GCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTG 5184
              |||||||||||||||||||||||| ||||||  ||||||| |||||||||||||||||||
Db       2208 GCAGGACTTAGCCCATACGAAATGCTATATGGATAGCCCTTCCTAACCAATGACCTTGTG 2267

Qy       5185 CTTGACCCA-AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAA 5243
              ||||||  | ||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db       2268 CTTGACTGAGAGAGAGCCAACTTAGTTGCAGACATCACCTCCTTATCCAAATATCAACAA 2327

Qy       5244 GTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCW 5303
              |||||||||||||||||||| ||| |||| |||||||||| || ||||||||||||||:
Db       2328 GTTCTTAAAACATTACAAGGAGCCTGTCCCCGAGAAGAGGGGAAGGAACTATTCCACCCT 2387

Qy       5304 WGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGG 5363
              |||||||||||||||||||||||||||:|||||||||  |   | ||||||| |||||||
Db       2388 GGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCTCATTGCCTAGATATATCCTGG 2447

Qy       5364 GAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAG 5423
              |||||||||||||||||||||||||:||||||||| || || ||||||||||||||||||
Db       2448 GAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGCAGTAAAAGTGGCTGGAGTGGAG 2507

Qy       5424 TCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCA 5483
              |||||||||||||||| |||||||| |||||||| | ||||||||||||||||||||||
Db       2508 TCTTGGATACATCACACTCGAGTCAAACCCTGGATATTACCAAAGGAACCTGAAAATCCA 2567

Qy       5484 GGAGACAA 5491
              ||||||||
Db       2568 GGAGACAA 2575


RESULT 80
AF123882
LOCUS       AF123882               2477 bp    mRNA    linear   VRL 11-AUG-1999
DEFINITION  Multiple sclerosis associated retrovirus element pol polyprotein
            (pol), integrase (pol) and envelope protein (env) pseudogenes,
            partial sequence.
ACCESSION   AF123882
VERSION     AF123882.1  GI:5726239
KEYWORDS    .
SOURCE      Multiple sclerosis associated retrovirus element
  ORGANISM  Multiple sclerosis associated retrovirus element
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 2477)
  AUTHORS   Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F.,
            Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E.,
            Mallet,F., Mandrand,B. and Perron,H.
  TITLE     Molecular cloning and characterization of MSRV-related sequences
            associated with retrovirus-like particles
  JOURNAL   Virology 260 (1), 1-9 (1999)
   PUBMED   10405350
REFERENCE   2  (bases 1 to 2477)
  AUTHORS   Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F.,
            Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E.,
            Mallet,F., Mandrand,B. and Perron,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1999) UMR103, CNRS/Biomerieux, 46 Allee D'Italie,
            Lyon Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .2477
                     /organism="Multiple sclerosis associated retrovirus
                     element"
                     /mol_type="mRNA"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:89382"
                     /clone="87-23"
     gene            <1. .1970
                     /gene="pol"
                     /note="pol polyprotein"
                     /pseudo
     misc_feature    769. .1970
                     /gene="pol"
                     /note="integrase"
                     /pseudo
     gene            1998. .>2477
                     /gene="env"
                     /note="envelope polyprotein"
                     /pseudo

  Query Match             28.2%;  Score 2109.4;  DB 1;  Length 2477;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches 2234;  Conservative   76;  Mismatches  160;  Indels   12;  Gaps    7;

Qy       3585 GGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGT 3644
              ||||||||||::|||||||||||||||| |||||||||||:|||||| |||||||
Db          2 GGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGGAAGAGATTTTACTGT 61

Qy       3645 GTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGA 3704
              |||||||||||| ||:|||||::|||||||| ||||||||||||||||||||||||||||
Db         62 GTGGACATCTCACGATGTGAACGGCATACTGACTGCTAAAGGAGACTTGTGGCTGTCAGA 121

Qy       3705 CAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCAC 3764
              ||||:|||||||||||||:|||||||||||||||||||:|||||||||||||:|||| |||
Db        122 CAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGCTGCGACTGCGCAT 181

Qy       3765 TTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAA 3824
              |||||||||||||||||||:| |||||||||| |||||||||||||||||:||:||||
Db        182 TTGTGCAACTCTTAACCCAGTCCCATTTCTTCCGGACAATGAAGAAAAGATAGAACATAA 241

Qy       3825 CTGTCAACAARTAATTTCTCAAACCTATGCCACTCGA-GGGGACCTTYTAGARGTTCCYT 3883
              ||||||||||:|||| |||||||||||  || ||||| ||||||| |:||||:|||||:|
Db        242 CTGTCAACAAGTAATTGCTCAAACCTACACCGCTCGAGGGGGACCCTTTAGAGGTTCCCT 301

Qy       3884 TGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTT 3943
              ||||||||||:|||| ||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTT 360

Qy       3944 CGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACT 4003
              | |||||: |||||||||||||||||||||||||||:|||||||||||| |||||||
Db        361 CAAAAAGCAGGGTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCTCCTCACT 420

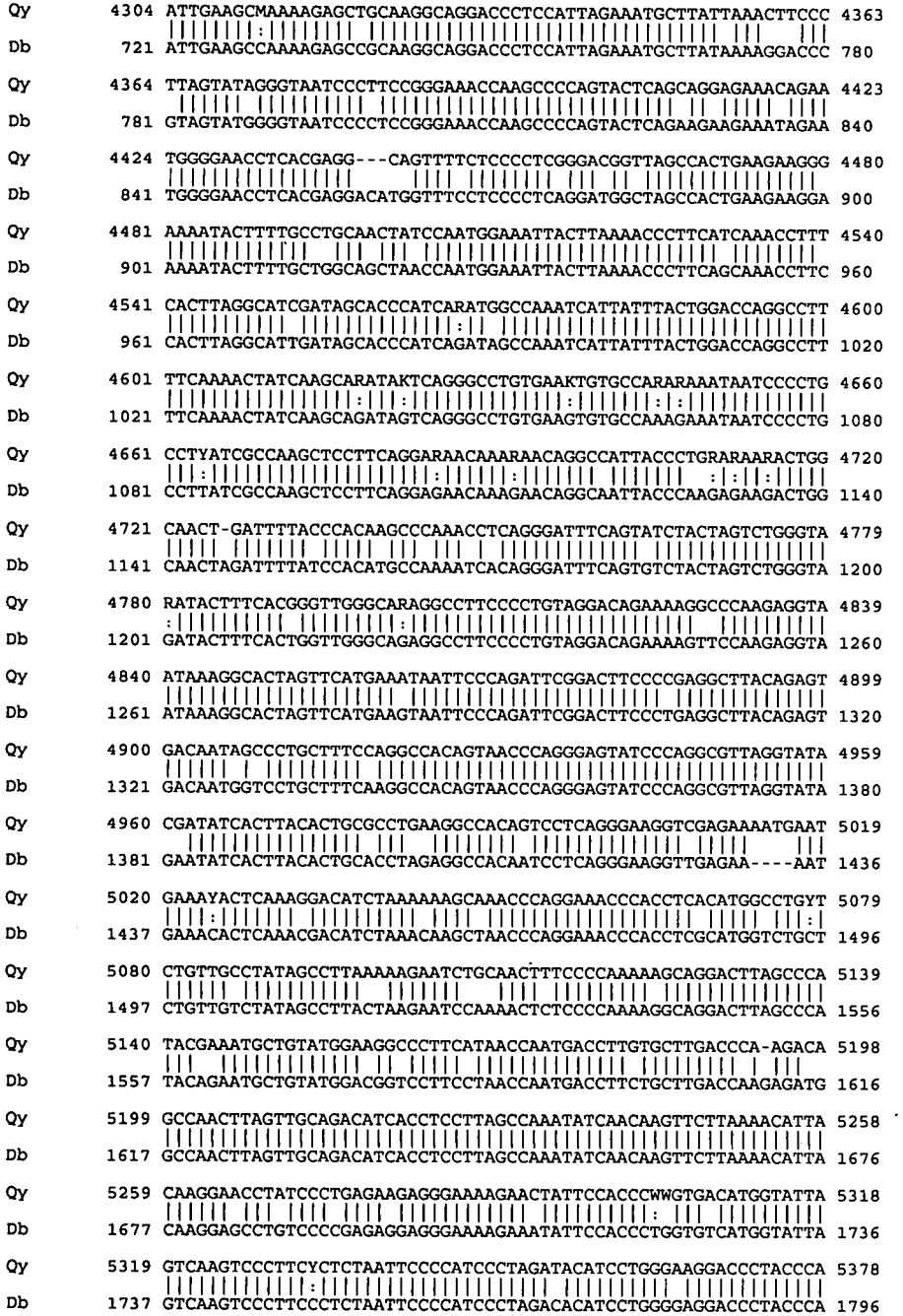
Qy       4004 CCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGA 4063
              || |||||||||||:||||:|||||||||||||| |:|||:| |||||||||||||||||
Db        421 CCGGGAACTAGTGCTCAGCTGGCAGAACTAATAGTCCTCACTCGGGCACTAGAATTAGGA 480

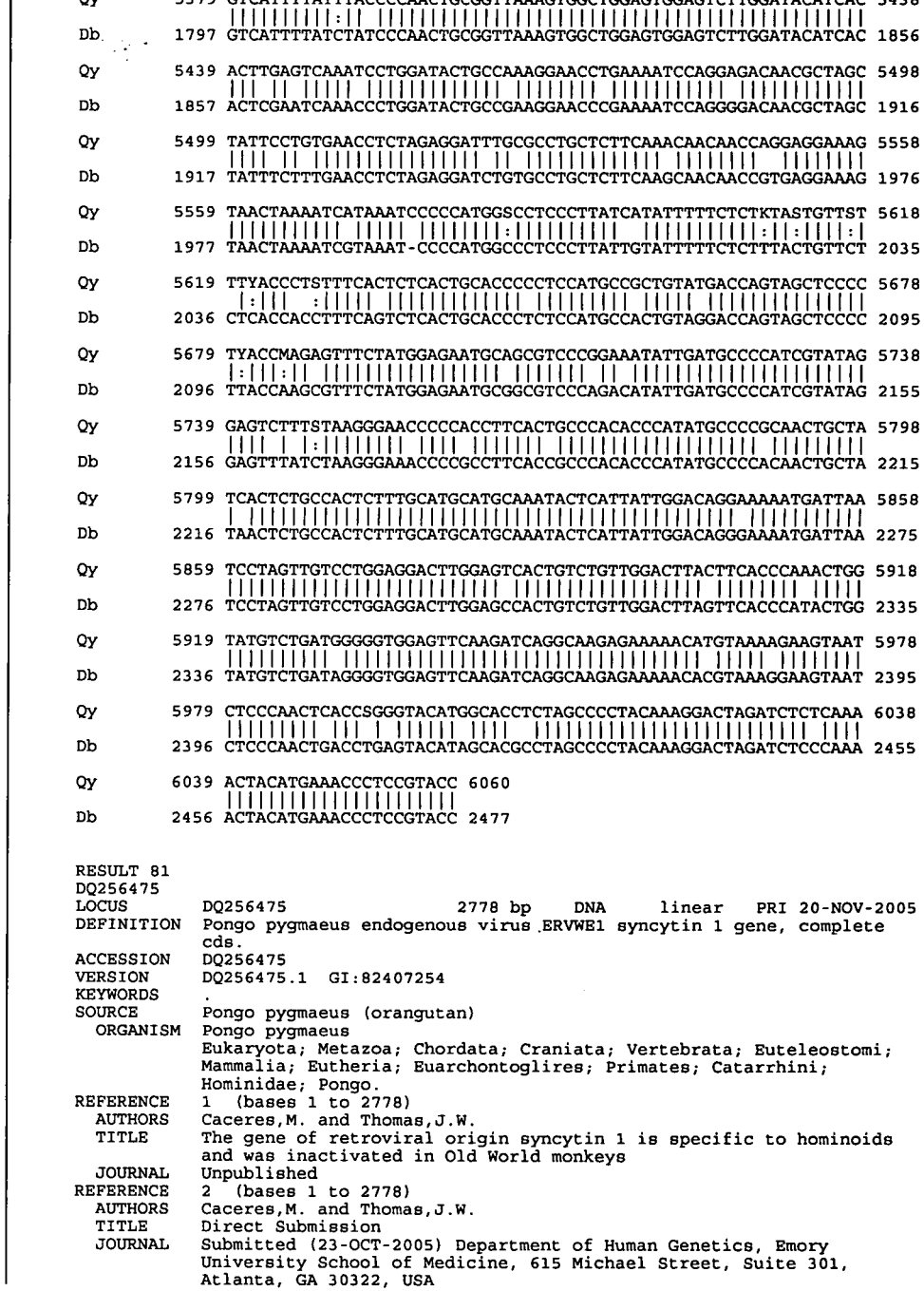
Qy       4064 GAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCC 4123
              ||||:|||||||:|||||||||||||||||:|:||||||:|||||||| |||||||||
Db        481 GAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTCCATGCC 540

Qy       4124 CATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACAT 4183
              ||||::|||||||:|:|||| |||||||||||| |||:|||:|||||||||||:||||
Db        541 CATGCAGCAATATGGAGAGAAGGGGAATTCCTAATTTCTGAGGGAACACCTATCAAACAT 600

Qy       4184 CAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTA 4243
              |||||||||||||||:|||||||:| ||:|||||| ||||||:||| ||| :|||||||
Db        601 CAGGAAGCCATTAGGAGATTATTATTCGCTGTACAGCAACCTAAAGAAGTGACAGTCTTA 660

Qy       4244 CACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKAT 4303
              ||||||:|||||||||| |||||||||:||||||||:|||:|||: |||||| :||
Db        661 CACTGCCGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGAGGAT 720
```





```
RESULT 81
DQ256475
LOCUS       DQ256475               2778 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Pongo pygmaeus endogenous virus ERVWE1 syncytin 1 gene, complete
            cds.
ACCESSION   DQ256475
VERSION     DQ256475.1  GI:82407254
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 2778)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2778)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
```

```
FEATURES             Location/Qualifiers
     source          1. .2778
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9600"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     CDS             197. .1813
                     /note="envelope glycoprotein"
                     /codon_start=1
                     /product="syncytin 1"
                     /protein_id="ABB73025.1"
                     /db_xref="GI:82407255"
                     /translation="MALPYHIFLFXVLLPSFTLTAPPPCRCMTSSSPYQEFVWRMHHP
                     GNIDAPSYRSFSKGTPTFTAHTHMPRNCYNSATLCMHGNTHYWTGKMINPSCPGGLGV
                     TVCWTYFTHTGMSDGGGVQDEAREKHVKEVISQLTQVHSTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWMCLPLAFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTNTTNSQCIRWVTPPTQVVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVVSKPRNKRVPILPFVMAAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSIAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNIGPWGLFSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYHRPLDWPASP
                     RSDVNDIKGTPPEEISTAQPLLRPNSAGSS"
     LTR             1872. .2651

  Query Match             28.0%;  Score 2100.7;  DB 1;  Length 2778;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches 2117;  Conservative   36;  Mismatches   45;  Indels    1;  Gaps    1;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
             |||||:|||||||| |||||||||||||||||||||||||||||||||||||||||||| |
Db          1 TTTATCTACCCCAACCGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
             ||||||| |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         61 AGTCAAACCCTGGATACTGCCAAAGGAACCCGAAAATCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
             |||||||||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db        121 CTGTGAACCTCTAGAGGATTTGTGCCTGCTCTTCAAACGACAACCAGGAGGAAAGTAACT 180

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
             |||||||||||| |||||||:|||||||||||||||||||||:|||||||:| |:||
Db        181 AAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTASTGTTCTCTTAC 239

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        240 CCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACC 299

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
             :|||||| |||||||||||||| | ||||||||||||||||||||||||||||||||||
Db        300 AAGAGTTTGTATGGAGAATGCATCATCCCGGAAATATTGATGCCCCATCGTATAGGAGTT 359

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             |||:|||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        360 TTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACT 419

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        420 CTGCCACTCTTTGCATGCATGGAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 479

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        480 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCATACTGGTATGT 539

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        540 CTGATGGGGGTGGAGTTCAAGATGAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 599

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| |||: ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        600 AACTGACCCAGGTACATAGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 659

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 719

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||| |||||||||||  |||||
Db        720 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGGCCTTCA 779

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 839

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             | |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        840 CTACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACC 899

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        900 TCACCTGTGTAAAATTTAGCAATACTACAAACACAACCAACTCCCAATGCATCAGGTGGG 959

Qy       6344 TAACTCCTCCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        960 TAACTCCTCCCCACACAAGTAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 1019

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1079

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:|||:|||||||||||||||||||||||||||||| ||||||| |||||||||||| |
Db       1080 CCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTATCTAAGCCCCGTA 1139

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db       1140 ACAAAAGAGTACCCATTCTTCCTTTTGTTATGGCAGCAGGAGTGCTAGGTGCACTAGGTA 1199

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CTGGCATTGGTGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1259

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCA 1319

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:|||||||| ||:|||||||||||
Db       1320 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCGAAAGAGGGG 1379

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1380 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATTGTCA 1439

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             |||||||||||:||||||||||:||||||||||||||||||:||||||:|||||||||||||
Db       1440 CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACA 1499

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||| ||||||:|||||||||||||||||||||||||||||||||||||
Db       1500 TTGGACCCTGGGGCCTCTTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1559

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||||||| ||||||||||||||||||||:|||||||||||||||||||
Db       1560 TAGCAGCTATAATATTGCTACTTCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACT 1619

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1620 TTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 1679

Qy       7064 CTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACA 7123
             ||||||||||| |||||||||||| ||||||:|||||||||||||||||||||:||||
Db       1680 CTAAGATCTACCACAGACCCCTGGACTGGCCTGCTAGCCCACGATCTGATGTTAAYGACA 1739

Qy       7124 TCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAG 7183
```

```
                ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        1740 TCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCTCTACTACGCCCCAATTCAG 1799

Qy        7184 CAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTT 7243
                |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        1800 CAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTT 1859

Qy        7244 GAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAA 7303
                |||||||||||||||||||||||||||||||||||||||||||  |:||||||||:|||
Db        1860 GAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCTACTAAGAATCCCTAA 1919

Qy        7304 GCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACA 7363
                ||||||:|||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        1920 GCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACA 1979

Qy        7364 CCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAAT 7423
                |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        1980 CCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAAT 2039

Qy        7424 AGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 7483
                ||||||||||:|||||||:||||||||||||||||||||||||||||||||||||||||
Db        2040 AGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 2099

Qy        7484 AGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTG 7543
                | :||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        2100 GGCCTTCGAGCCAGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTG 2159

Qy        7544 TTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                |||||||||||||||||||||||||||||||:||||:
Db        2160 TTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 2198


RESULT 82
DQ256474
LOCUS       DQ256474               2732 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Gorilla gorilla endogenous virus ERVWE1 syncytin 1 gene, complete
            cds.
ACCESSION   DQ256474
VERSION     DQ256474.1  GI:82407252
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (bases 1 to 2732)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2732)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2732
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9593"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     CDS             197. .1813
                     /note="envelope glycoprotein"
                     /codon_start=1
                     /product="syncytin 1"
                     /protein_id="ABB73024.1"
                     /db_xref="GI:82407253"
                     /translation="MALPYHILLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMRRP
                     GNIDAPSXRSFSKGTPTFTAHTHMPRNCYNSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTHTGMSDGGGVQDQAREKHVKEVISQLTRVHSTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLDFRPYVSIPVPEEWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTIDTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVVSKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKCTPPEEISTAQPLLRPNSAGSS"
     LTR             1872. .2605

  Query Match             27.8%;  Score 2080.3;  DB 1;  Length 2732;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches 2086;  Conservative   39;  Mismatches   31;  Indels    1;  Gaps    1;

Qy        5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
                |||||:|||||||||| | ||||||||||||||||||||||||||||||||||||| |
Db           1 TTTATCTACCCCAACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCG 60

Qy        5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
                ||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          61 AGTCAAACCCTGGATACTGCCAAAGGAACCTGAAAATCCAGCAGACAACGCTAGCTATTC 120

Qy        5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 180

Qy        5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
                |||||||||||| |||||||:|||||||||||||| |||||||:||:||||:|||:||
Db         181 AAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTCTTCTCTTTACTGTTCTTTTAC 239

Qy        5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
                |||:|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db         240 CCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACC 299

Qy        5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
                :|||||||||||||||||||| |||||| |||||||||||||||||||:||||||||
Db         300 AAGAGTTTCTATGGAGAATGCGGCGTCCTGGAAATATTGATGCCCCATCGYATAGGAGTT 359

Qy        5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
                |||:|||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         360 TTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACT 419

Qy        5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         420 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 479

Qy        5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
                |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         480 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCATACTGGTATGT 539

Qy        5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         540 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 599

Qy        5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
                |||| |||:|||||||| ||||||||||||||||||||||||||||||||||||||||
Db         600 AACTGACCCGGGTACATAGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 659

Qy        6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         660 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 719

Qy        6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
                |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         720 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGGACTTCA 779

Qy        6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
                :|||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         780 GGCCATATGTTTCAATCCCTGTACCTGAAGAATGGAACAACTTCAGCACAGAAATAAACA 839

Qy        6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
                ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db         840 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACC 899

Qy        6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
```

```
Db        900 TCACCTGTGTAAAATTTAGCAATACTATAGACACAACCAACTCCCAATGCATCAGGTGGG 959

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403

Db        960 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTCGTCTGTGGTACCT 1019

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463

Db       1020 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1079

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523

Db       1080 CTCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTATCTAAGCCCCGCA 1139

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583

Db       1140 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCRGGAGTGCTAGGTGCACTAGGTA 1199

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643

Db       1200 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1259

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703

Db       1260 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 1319

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763

Db       1320 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGG 1379

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823

Db       1380 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 1439

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883

Db       1440 CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACA 1499

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943

Db       1500 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1559

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003

Db       1560 TTGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACT 1619

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063

Db       1620 TTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 1679

Qy       7064 CTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACA 7123

Db       1680 CTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACA 1739

Qy       7124 TCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAG 7183

Db       1740 TCAAATGCACCCCTCCCGAGGAAATCTCAACTGCACAACCWCTACTACGCCCCAATTCAG 1799

Qy       7184 CAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTT 7243

Db       1800 CAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTT 1859

Qy       7244 GAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAA 7303

Db       1860 GAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAA 1919

Qy       7304 GCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACA 7363

Db       1920 GCCTAGCTGGGAAGGTGACCACATCCATCTTTAAACACAGGGCTTGCAACTTAGCTCACA 1979

Qy       7364 CCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAAT 7423

Db       1980 CCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAAT 2039

Qy       7424 AGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 7483

Db       2040 AGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCA 2099

Qy       7484 AGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCT 7540

Db       2100 GGCCTTCGAGCCGGTAACRGCAACCCCCTTTGGGTCCCCTCYCTTTGTATGGGAGCT 2156
```

```
RESULT 83
DQ256476
LOCUS       DQ256476               2764 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Symphalangus syndactylus endogenous virus ERVWE1 syncytin 1 gene,
            complete cds.
ACCESSION   DQ256476
VERSION     DQ256476.1  GI:82407256
KEYWORDS    .
SOURCE      Symphalangus syndactylus (siamang)
  ORGANISM  Symphalangus syndactylus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Symphalangus.
REFERENCE   1  (bases 1 to 2764)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2764)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2764
                     /organism="Symphalangus syndactylus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9590"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     CDS             197. .1813
                     /note="envelope glycoprotein"
                     /codon_start=1
                     /product="syncytin 1"
                     /protein_id="ABB73026.1"
                     /db_xref="GI:82407257"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQHP
                     GNIDAPSYRSFSKGTPTFTAHTHMPRNCYNSATLCMHANTQYWTGKMINPSCPGGLGV
                     TVCWTYFTHTGMSDGGGVQNQARKEKHVKEVISQLTQVHSTSSPYKGLDLSKLHETLHT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWMCLPLDFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTDTTNSQCIRWVTPPTRIFCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVVSKPRNKRVPILPFVMGAGVLGALG
                     TGIGGITTSTQFYYKLSRELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAIKLQMEPKMESKTKTYHRSLDWPVSP
                     RSDVNDIKGIPPEEISTAQPLLRPNSAGSS"
     LTR             1872. .2637

  Query Match             27.7%;  Score 2074.3;  DB 1;  Length 2764;
  Best Local Similarity   95.5%;  Pred. No. 0;
  Matches 2101;  Conservative   35;  Mismatches   62;  Indels    1;  Gaps    1;
```

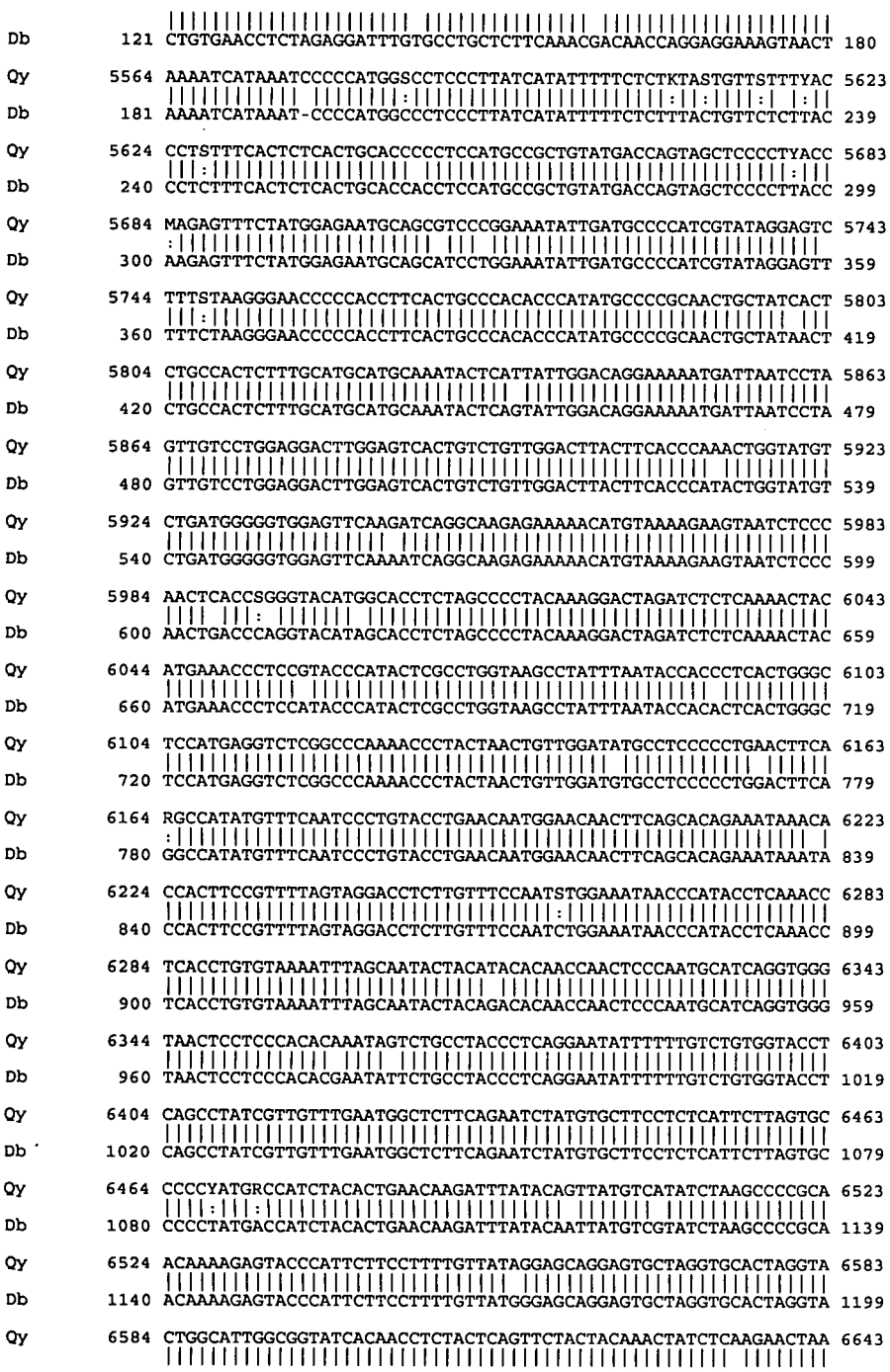
```
Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443

Db          1 TTTATCTACCCCAACCGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503

Db         61 AGTCAAACCCTGGATACCGCCAAAGGAACCCGAAAATCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
```




RESULT 84
BD196247
LOCUS BD196247 2585 bp DNA linear PAT 17-JUL-2003
DEFINITION Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.
ACCESSION BD196247

```
VERSION     BD196247.1  GI:33006017
KEYWORDS    JP 2002512530-A/13.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2585)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 13 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/13
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2585
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2585
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             27.2%;  Score 2036.1;  DB 1;  Length 2585;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches 2233;  Conservative  129;  Mismatches  213;  Indels   19;  Gaps   14;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
Db         61 ACTTGGACACTCTTGTCCTTTGGTATGTGGATGATCTACTTTTAGCCACCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
Db        121 CCTTGTGCCATCAAGCCAACCAAGTGCTCTTAAACTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db        181 TTTCCAAACCAGAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db        301 CCAAAACCCTGAAGCAATTAAGAGGGTTCCTTGGCATAAAAGGCTGCTGTTGAATATGGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db        361 TT-CCCAGGTACAATGAAATAGCCAGGCCATTATACACACTAATTACGGGAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db        420 GCCAATACCCATTTAGTAGAATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAG 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db        480 AAGGCCCT--AATCCAAGCCCCAGTGTTAAGCTTGCCAATGGAGCAAGACTTTTCTTTAT 537

Qy       3448 AYRTCACAG--AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
Db        538 ATGTCACAGAAAAAAAAACAGGAATAGCTCTAGAAGTCCTTACACAGGTCCGAGGGACCA 597

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
Db        598 GCTTACAACACATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGACTC 657

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
Db        658 ATTGTTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGCATCTGAAGCAGTTAAAATGATAC 717

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
Db        718 AGGGAAGANATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAG 777

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
Db        778 GAGAC-TGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATCACTTGAANGGC 836

Qy       3746 CAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATG 3805
Db        837 CAGTGCTGCCACTGTGCACTTGTGCAACTCTTAACCCACCCACATTTCTTCCAGACAATG 896

Qy       3806 AAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGG 3865
Db        897 AAGAAAAGATAGAACATAACTGTCAACAAGTGATTGTTCAAACCTACACCGCTCGAAGGG 956

Qy       3866 ACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCC 3925
Db        957 ACCTTCTAGAGGTTCCCTTGACTGATCCTGAGC-TCAACTTCTATACTGATGGAAGTTCC 1015

Qy       3926 TTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTT 3985
Db        1016 TTTTGTAGAAAAAGGACTTCGAAAGGCGGGTATGCAGTGGCCAGTGATAATGGAATACTT 1075

Qy       3986 GAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYT 4045
Db        1076 GAAAGTAATCCCTTCACTCCAGAAACTAGCATTCAGCTGGCAGAATTAATAGCCTTCACT 1135

Qy       4046 KGGGCACTAGAA-TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGC 4104
Db        1136 TGGGCATTAGAACACAGGAGAAGGAAAAGGAGTAAATATATATACAGACTCCAAGTATGC 1195

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
Db        1196 TTACTTAGTCCTCCATGCCCATGCAGCAATATAGAGAGAAAGCGAATTCCTAACTTCTGA 1255

Qy       4165 GRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACC 4224
Db        1256 GGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTACTGGCTGTACAGAAACC 1315

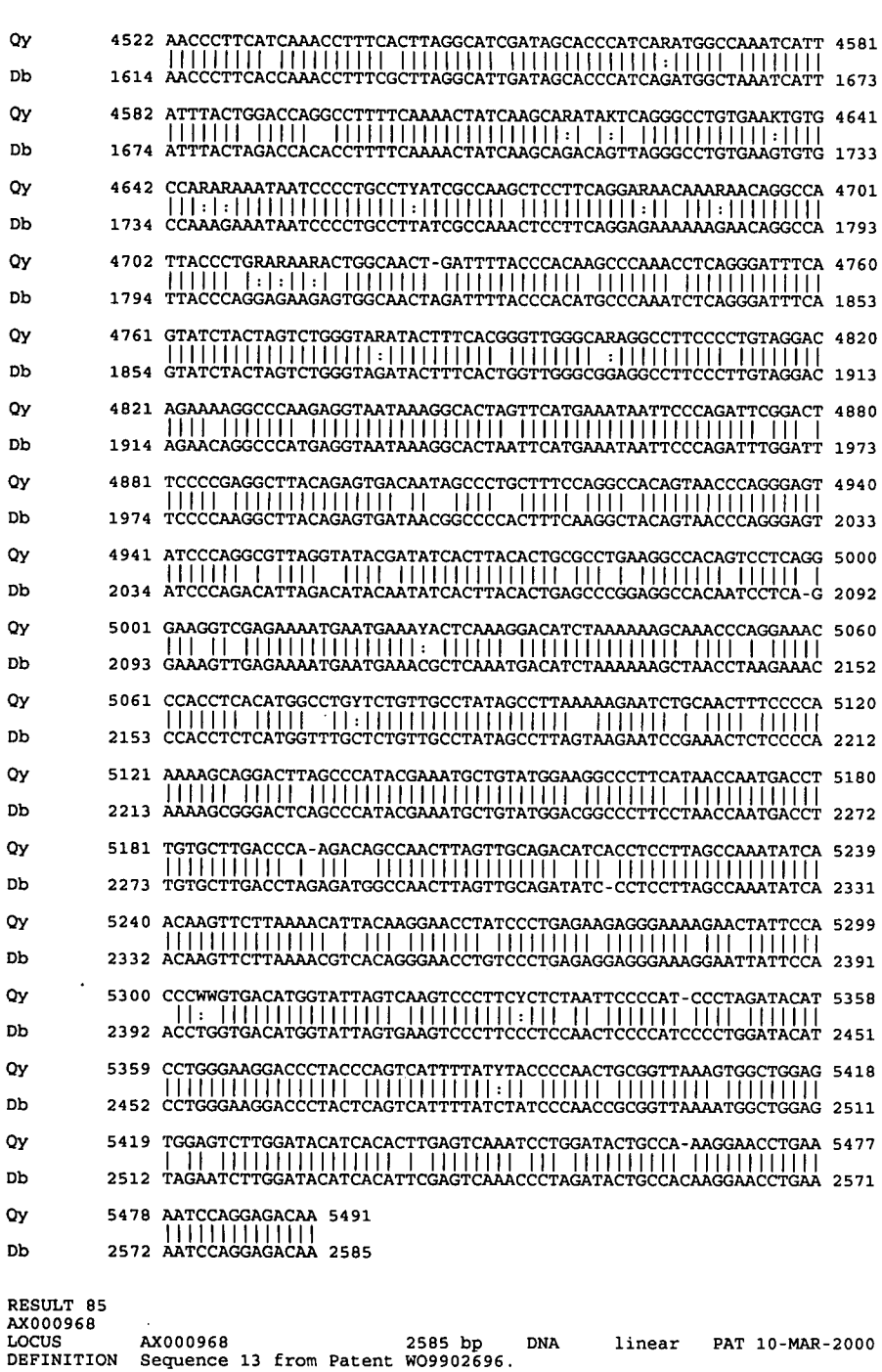
Qy       4225 TARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASA 4284
Db        1316 T--AGAGGTGGCAGTCTTACATGGCCGAGATCATCAGAAAGGAAAAGAAAGGGAAATAGA 1373

Qy       4285 AGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGA 4344
Db        1374 AGGGAACTGCCAAGTGGATATTGAAGCCAAAAGAGCTGCAAGGCGGGACCCTCCATTAGA 1433

Qy       4345 AATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTA 4404
Db        1434 AATGCTTATAGAAGGACCCCTAGTACAGGGCAATCCCCTTCAGGAAACCAAGCCCCAATA 1493

Qy       4405 CTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACG 4461
Db        1494 CTCAGCAGAAGAAATGGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATG 1553

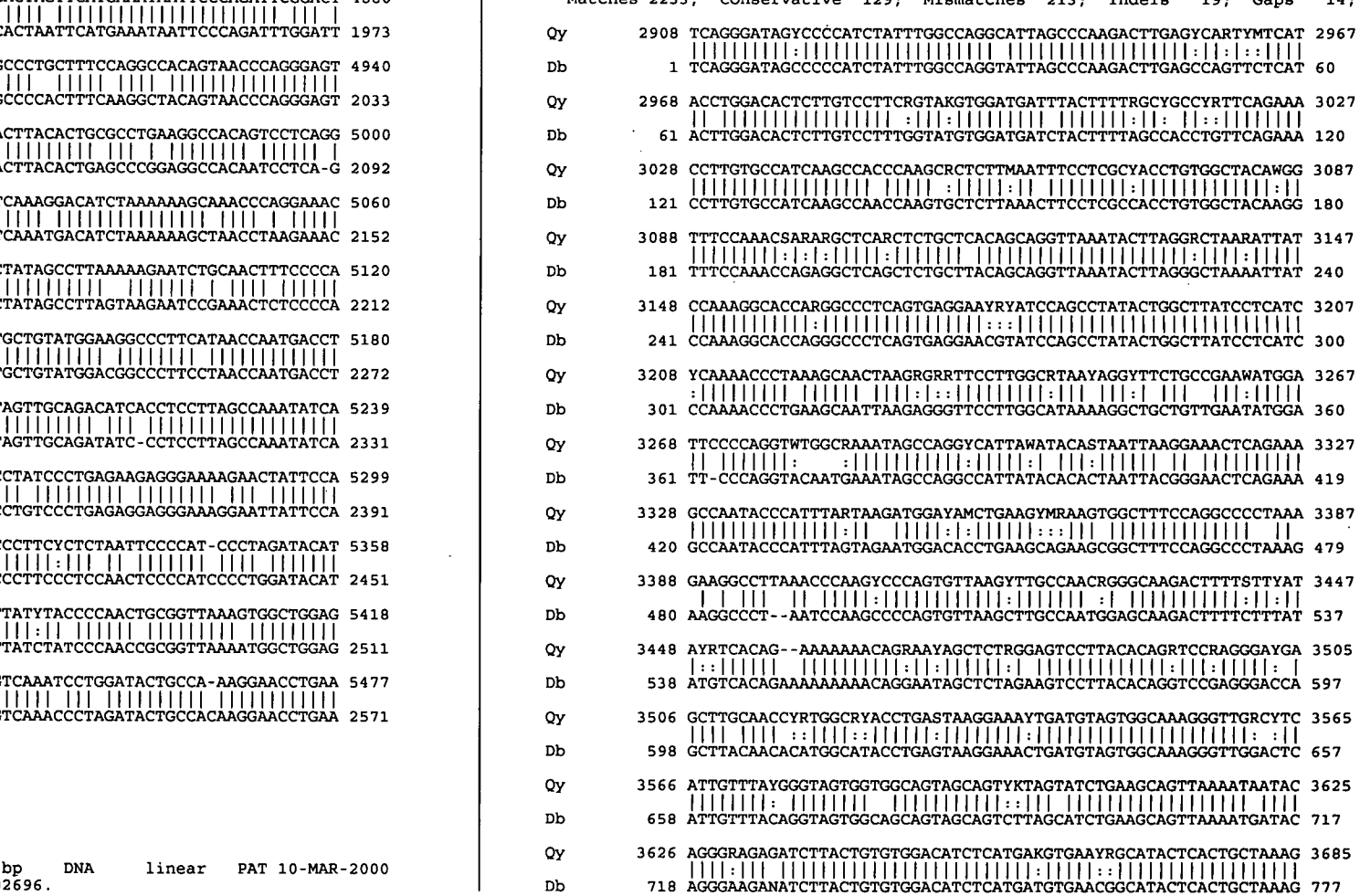
Qy       4462 GTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAA 4521
Db        1554 GCTAGCCACCAAAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAA 1613
```

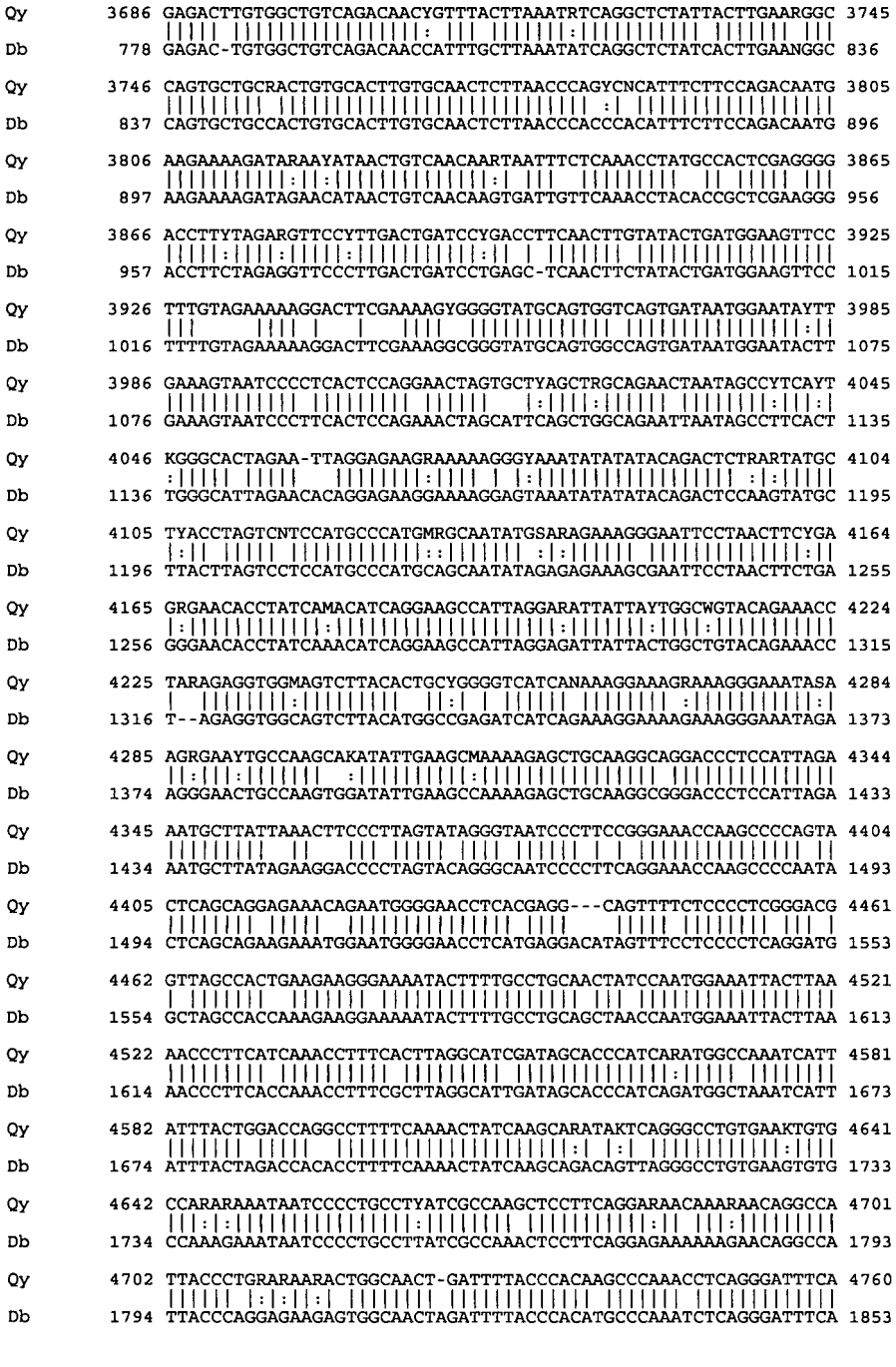


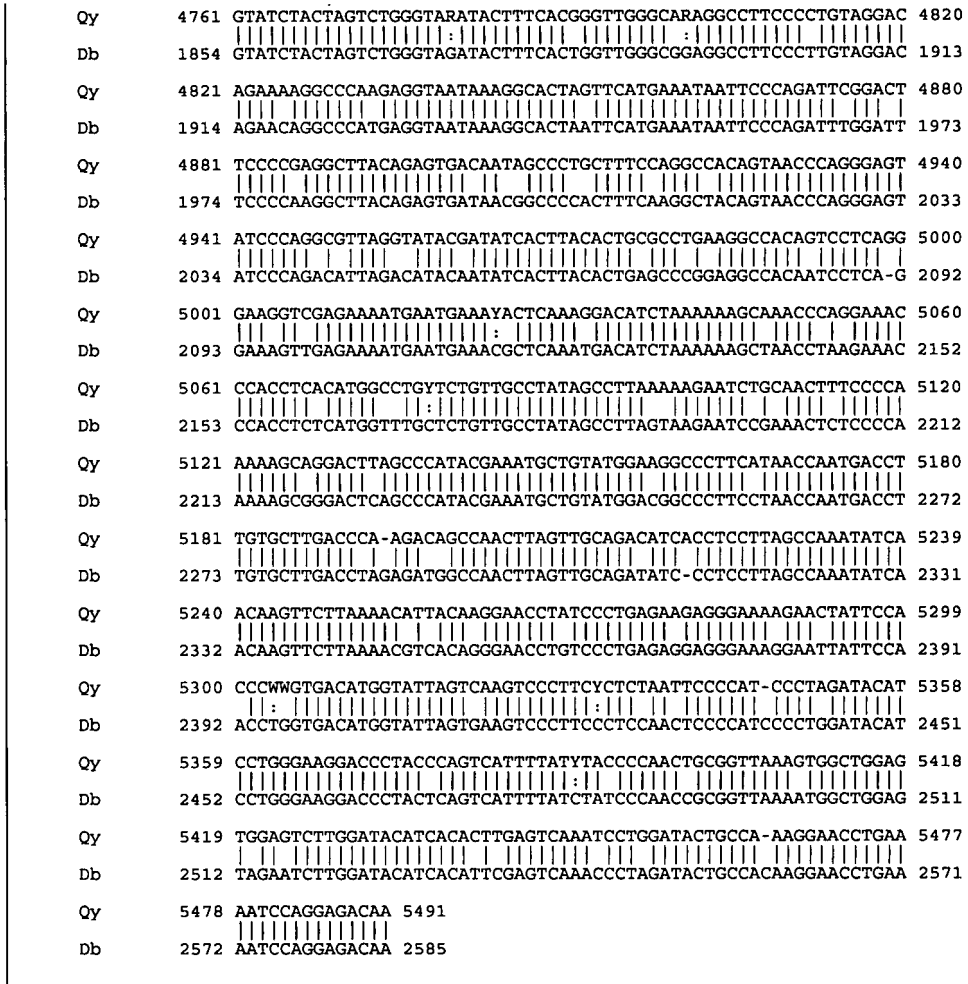
RESULT 85
AX000968
LOCUS AX000968 2585 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 13 from Patent WO9902696.

ACCESSION AX000968
VERSION AX000968.1 GI:7241210
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 2585)
AUTHORS Beseme,F. and Blond,J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES OR WITH PREGNANCY DISORDERS
JOURNAL Patent: WO 9902696-A 13 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES Location/Qualifiers
source 1. .2585
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 27.2%; Score 2036.1; DB 1; Length 2585;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2233; Conservative 129; Mismatches 213; Indels 19; Gaps 14;







```
RESULT 86
BD221827
LOCUS       BD221827               2055 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221827
VERSION     BD221827.1  GI:33031597
KEYWORDS    JP 2002518051-A/22.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2055)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 22 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
```

```
PN   JP 2002518051-A/22
PD   25-JUN-2002
PF   23-JUN-1999 JP 2000556036
PR   23-JUN-1998 FR     98/07920
PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
A61P25/00,
PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
sequence and deduced protein sequence family with  CC
human endogenous
CC   retroviral motifs, and their uses
FH   Key             Location/Qualifiers
FT   CDS             (1). .(2055).
FEATURES             Location/Qualifiers
     source          1. .2055
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             27.1%;  Score 2032.5;  DB 1;  Length 2055;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 2025;  Conservative   30;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db         61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
             |||||||||||||||||||||||||| ||||||||:||||||||||||||||||||||||:
Db        361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
             ||:||||:|||:|||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
             ||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
             |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
             |||||||||||||||||||||||||||||||:|||||||||||||||||||:|||||||
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
             ||||||||||||||||||||||||||:|||||||||:|||||||||||||||:||||||
Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
             :|||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859
```

```
Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245
             ||||||||||||||||
Db       2040 TAGGTTTTCCTGTTGA 2055


RESULT 87
AR699776
LOCUS       AR699776               2055 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 22 from patent US 6919438.
ACCESSION   AR699776
VERSION     AR699776.1  GI:75205567
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2055)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 22 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .2055
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             27.1%;  Score 2032.5;  DB 1;  Length 2055;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 2025;  Conservative   30;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||::|||||||
Db         61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
             |||||||||||||||||||||||| ||||||||:|||||||||||||||||||||||||:
Db        361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419
```

```
Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
             ||:||||:|||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
             ||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079

Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
```

```
                |||||||||||||||||||||||||||||||:||||||||||||||||||||:||||||
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
                |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
                |||||||||||||||||||||||||:|||||||||:||||||||||||||||||:||||||
Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
                :||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
                |||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
                ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
                ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245
                ||||||||||||||||
Db       2040 TAGGTTTTCCTGTTGA 2055
```

```
RESULT 88
AX007999
LOCUS       AX007999               2055 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 22 from Patent WO9967395.
ACCESSION   AX007999
VERSION     AX007999.1  GI:9995696
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 22 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .2055
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1. .87
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07392.1"
                     /db_xref="GI:9995697"
                     /translation="PKTANLVADITSLAKYQQVLKTLQGTYP"

  Query Match             27.1%;  Score 2032.5;  DB 1;  Length 2055;
  Best Local Similarity   98.5%;  Pred. No. 0;
```

```
  Matches 2025;  Conservative   30;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5190 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 5249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAAGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTT 60

Qy       5250 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGAC 5309
                ||||||||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db          61 AAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGAC 120

Qy       5310 ATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 5369
                |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         121 ATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGA 180

Qy       5370 CCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 5429
                |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         181 CCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGG 240

Qy       5430 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 5489
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGAC 300

Qy       5490 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 5549
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCA 360

Qy       5550 GGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTK 5609
                |||||||||||||||||||||||||| ||||||||:||||||||||||||||||||||:
Db         361 GGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTT 419

Qy       5610 TASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 5669
                ||:||||:|||:|||||:||||||||||||||||||||||||||||||||||||||||||
Db         420 TACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAG 479

Qy       5670 TAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 5729
                ||||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db         480 TAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCC 539

Qy       5730 ATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 5789
                ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         540 ATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCG 599

Qy       5790 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 5849
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         600 CAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA 659

Qy       5850 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 5909
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         660 AATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAC 719

Qy       5910 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 5969
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         720 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAA 779

Qy       5970 AGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 6029
                ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         780 AGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGA 839

Qy       6030 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 6089
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         840 TCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATAC 899

Qy       6090 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 6149
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         900 CACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCT 959

Qy       6150 CCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 6209
                ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         960 CCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAG 1019

Qy       6210 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAAC 6269
                ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        1020 CACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAAC 1079
```

```
Qy       6270 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 6329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA 1139

Qy       6330 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 6389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTT 1199

Qy       6390 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 6449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCT 1259

Qy       6450 CTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCAT 6509
             ||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||||||
Db       1260 CTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCAT 1319

Qy       6510 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 6569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCT 1379

Qy       6570 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 6629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACT 1439

Qy       6630 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 6689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGA 1499

Qy       6690 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAAC 6749
             ||||||||||||||||||||||||||||||:||||||||||||||||||:||||||||||
Db       1500 TCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAAC 1559

Qy       6750 CGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 6809
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCA 1619

Qy       6810 ATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGA 6869
             ||||||||||||||||||||||||:|||||||||:||||||||||||||||||:||||||
Db       1620 ATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGA 1679

Qy       6870 RGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCC 6929
             :||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db       1680 GGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCC 1739

Qy       6930 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRA 6989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1740 CTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAA 1799

Qy       6990 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAA 7049
             ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db       1800 CCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAA 1859

Qy       7050 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1860 GATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATC 1919

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 1979

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1980 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACT 2039

Qy       7230 TAGGTTTTCCTGTTGA 7245
             ||||||||||||||||
Db       2040 TAGGTTTTCCTGTTGA 2055
```

```
RESULT 89
DQ256480
LOCUS       DQ256480               2346 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Cercopithecus hamlyni endogenous virus ERVWE1 syncytin 1
            pseudogene, partial sequence.
ACCESSION   DQ256480
VERSION     DQ256480.1  GI:82407261
KEYWORDS    .
SOURCE      Cercopithecus hamlyni (owl-faced monkey)
  ORGANISM  Cercopithecus hamlyni
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Cercopithecus.
REFERENCE   1  (bases 1 to 2346)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2346)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2346
                     /organism="Cercopithecus hamlyni"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9536"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2199
                     /gene="syncytin 1"
                     /pseudo
     LTR             1873. .2302

  Query Match            26.4%;  Score 1979.8;  DB 1;  Length 2346;
  Best Local Similarity  93.0%;  Pred. No. 0;
  Matches 2047;  Conservative   37;  Mismatches  114;  Indels    2;  Gaps    2;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
             |||||:|||||||||||||||||||||||||||||||||||| |||||||||||||||| |
Db          1 TTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCATGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
             ||||||| ||||||||| | |||||||||||||||| ||||||||||||||||||||||||
Db         61 AGTCAAACCCTGGATACAGTTAAAGGAACCTGAAAACCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACGACAACCAGGAGGAAAGTAACT 180

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
             |||||| ||||| |||||||||:|||||||||||||||||| |||:||:||||:| |:||
Db        181 AAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTTTTTTCTTTACTGTTCTCTTAC 239

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
             || :|||||||||||||||||||||||||||||||||||| ||||||||||||||||:|||
Db        240 CCCCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTACGACCAGTAGCTCCCCTTACC 299

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
             :|||||||||:|||||||||||||||| |||||||||||||||||| ||||||||||
Db        300 AAGAGTTTCTATRGAGAATGCAGCGTCCTGGAAATATTGATGCCCCATTGTATAGGAGTT 359

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||  | ||| |
Db        360 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAATGGTTATAATG 419

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||  |||||||||| ||||||||||||||||||||||||||||||
Db        420 CTGCCACTCTTTGCATGCAATCAAATACTCACTATTGGACAGGAAAAATGATTAATCCTG 479

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             ||||||:|||||||||||||| ||||||||||||||||||||||||| ||  ||||||
Db        480 GTTGTCYTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGT 539
```

```
Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             | ||||||||||||||||||:||||||||||||| |||||  ||||||||||||||||||
Db        540 CAGATGGGGGTGGAGTTCAAKATCAGGCAAGAGAGAAACACATAAAAGAAGTAATCTCCC 599

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             ||||  |||:|||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        600 AACTGACCCGGGTACATAGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 659

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCATTGGGC 719

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        720 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCA 779

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGCCATACATTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 839

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||||||||||||||| :|||||||||||  ||||||||||||
Db        840 CCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAACCTGGAAATAACCTGTACCTCAAACC 899

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||| ||||||| ||||| | | ||||||||||||||||||||||||||||
Db        900 TCACCTGTGTAAATTTTAGCAGTACTATAGATACAACCAACTCCCAATGCATCAGGTGGG 959

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        960 TAACTCCTCCCACACGAATAGTCTGCCTATCCTCAGGAATATTTTTTGTCTGTGGTACCT 1019

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1079

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             | ||:|||:|||||||||||||||||||||||||||| ||||||||||  ||||||||||
Db       1080 CACCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATACATAAGCCCCGCA 1139

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||||
Db       1140 ACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGGTGGACTAGGTA 1199

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1259

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db       1260 ATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 1319

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||| |:|||||| || ||:|||||||
Db       1320 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAATTGCTAACTGCCGAAAGAGGGG 1379

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Db       1380 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCA 1439

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             ||||||||||:|||||||   |||| ||||||||||:|||| |:|||||| |||||
Db       1440 GTGAGAAAGTTAAAGAAATTAAAGATTGAATACAACGTAGAGCAAAGGAGCTTCAAAACA 1499

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||  |||||||||||||||||||||  ||
Db       1500 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCACGGATTCTCCCCTTCTTAGGACTTC 1559

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||| |
Db       1560 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1619

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063
             |||||||||| |||||| |||||||:||||||||||||||||||| |||||||||||| ||
Db       1620 TTGTCTCTTCTAGAATCAAAGCTGTAAAACTACAAATGGAGCCCCAGATGCAGTCCATGA 1679

Qy       7064 C-TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
             | |||| ||||||||||||||||||||||||||||||:|||||| |||| ||||||||||
Db       1680 CTTAAGTTCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCATGATCCGATGTTAATGAC 1739

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
             |||||||||||||||||||||||||||||| |||||||||| |||||||||||| ||||
Db       1740 ATCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTACGCCCCAGTTCA 1799

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
             |||||||||||||||| ||: ||||||||||||||||||||||||||||||||||||||||
Db       1800 GCAGGAAGCAGTTAGAGCAGTCATCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 1859

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
             ||||| |||| ||| |||||||||||||| ||||||||||||    |:||||||||:||
Db       1860 TGAGAGGGGGCACTAAGAGACAGGACTAGTTGGATTTCCTAGGACAACTAAGAATCCCTA 1919

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
             || ||||:||||||||||||| ||| |||||||||||| |||||||||||||||:||||
Db       1920 AGGCTAGCTGGGAAGGTGACCGCATTCACCTTTAAACAGGGGGCTTGCAACTTAGCTCAC 1979

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
             ||| ||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||
Db       1980 ACCCGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGAAGGTAAAGAAA 2039

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
             ||||||||||:|||||:|||||||||||  |||||||||| ||||||||||||||||||||
Db       2040 TAGCCAATCATCTATTGCCTGAGAGCACAGTGGGAGGGACAAGGATCGGGATATAAACCC 2099

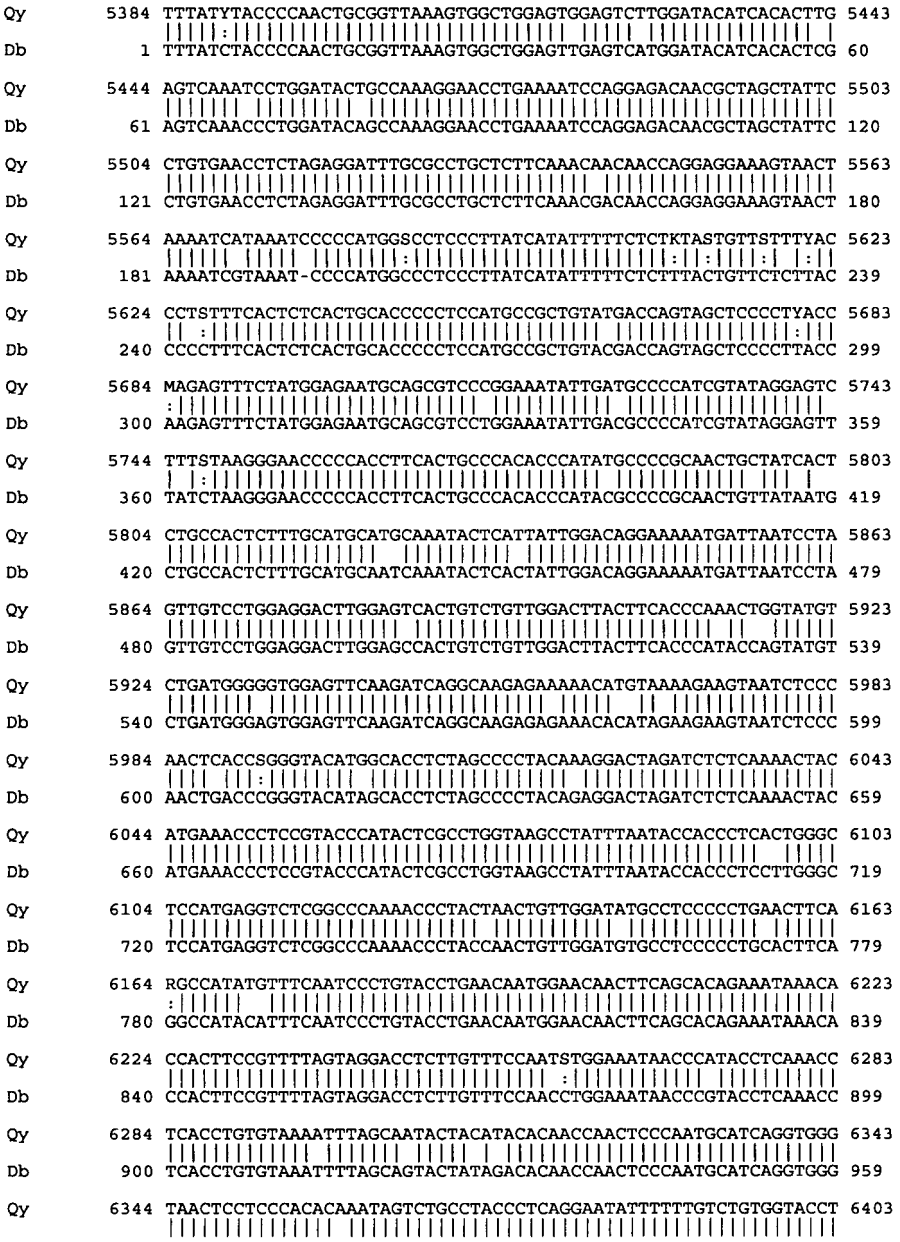
Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
             | |  ||| ||| |||||| ||||||||||||||||||||||||||||||||||||||||
Db       2100 AGGCATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 2159

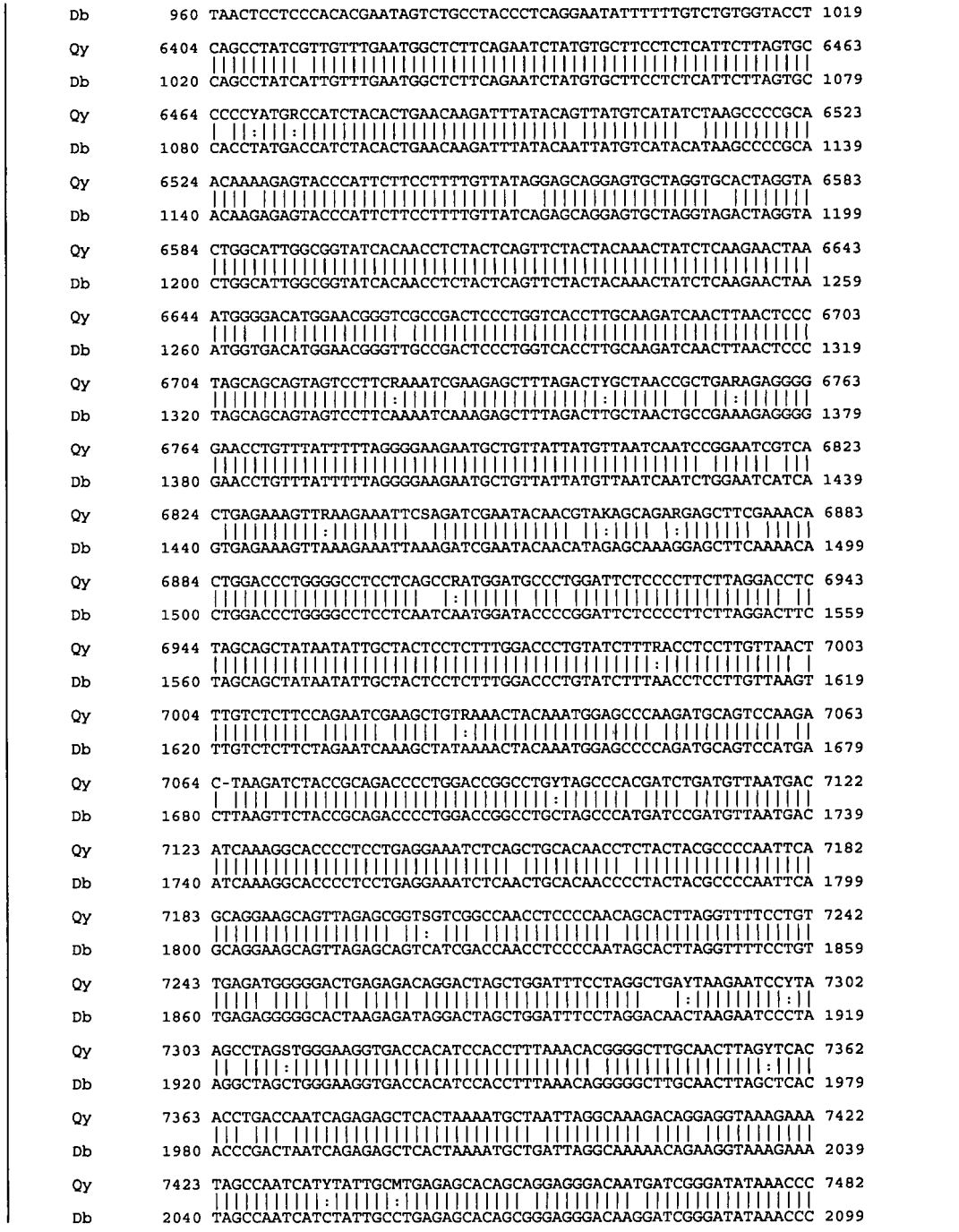
Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||  |||||||||||||||||||||||:||||:
Db       2160 GTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 2199
```

```
RESULT 90
DQ256481
LOCUS       DQ256481               2343 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Macaca mulatta endogenous virus ERVWE1 syncytin 1 pseudogene,
            partial sequence.
ACCESSION   DQ256481
VERSION     DQ256481.1  GI:82407262
KEYWORDS    .
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 2343)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2343)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2343
                     /organism="Macaca mulatta"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9544"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2199
                     /gene="syncytin 1"
```

/pseudo
LTR 1873. .2299

Query Match 26.4%; Score 1977.8; DB 1; Length 2343;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 2048; Conservative 34; Mismatches 116; Indels 2; Gaps 2;





```
Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
Db       2100 AGGCATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 2159

Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
Db       2160 GTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTACA 2199


RESULT 91
DQ256483
LOCUS       DQ256483               2341 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Papio hamadryas endogenous virus ERVWE1 syncytin 1 pseudogene,
            partial sequence.
ACCESSION   DQ256483
VERSION     DQ256483.1  GI:82407264
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 2341)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2341)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2341
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9557"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2197
                     /gene="syncytin 1"
                     /pseudo
     LTR             1871. .2297

  Query Match             26.2%;  Score 1964.8;  DB 1;  Length 2341;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches 2053;  Conservative   33;  Mismatches  110;  Indels    4;  Gaps    4;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
Db          1 TTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCATGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
Db         61 AGTCAAACCCTGGATACAGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
Db        121 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACGACAACCAGGAGGAAAGTAACT 180

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
Db        181 AAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTCTTAC 239

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
Db        240 CCCCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTACGACCAGTAGCTCCCCTTACC 299

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
Db        300 AAGAGTTTCTATGGAGAATGCAGCGTCCTGGAAATATTGATGCCCCATTGTATAGGAGTT 359

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
Db         360 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGTTATAATG 419

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
Db         420 CTGCCACTCTTTGCATGCAATCAAATACTCACTATTGGACAGGAAAAATGATTAATCCTA 479

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
Db         480 GTTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACCTGTATGT 539

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
Db         540 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACACATAGAAGAAGTAATCTCCC 599

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
Db         600 AACTGACCCGGGTACATAGCACCTCTAGCCCCTGCAGAGGACTAGATCTCTCAAAACTAC 659

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
Db         660 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCCTTGGGC 719

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
Db         720 TCCATGAGGTCTCGGCCTAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCA 779

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
Db         780 GGCC-TACATTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 838

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
Db         839 CCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAACCTGGAAATAACCCGTACCTCAAACC 898

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
Db         899 TCACCTGTGTAAATTTTAGCAGTACTATAGACACAACCAACTCCCAATGCATCAGGTGGG 958

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
Db         959 TAACTCCTCCCACATGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGCGGTACCT 1018

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
Db         1019 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1078

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
Db         1079 CACCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATACATAAGCCCCGCA 1138

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
Db         1139 ACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGGTGGACTAGGTA 1198

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
Db         1199 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1258

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
Db         1259 ATGGTGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 1318

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
Db         1319 TAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGCTAACTGCCGAAAGAGGGG 1378

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
Db         1379 GAACCTCTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCA 1438
```

```
Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
Db       1439 GTGAGAAAGTTAAAGAAATTAAAGATCGAATACAACGTAGAGCAAAGGAGCTTCAAAACA 1498

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
Db       1499 CTGGACCCTGGGGCCTCCTCAGTCGATGGATGCCCTGGATTCTCCCCTTCTTAGGACTTC 1558

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
Db       1559 TAGCAGCTATAATATTGCTACTCCTCCTTGGACCCTGTATCTTTAACCTCTTTGTTAAGT 1618

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063
Db       1619 TTGTCTCTTCTAGAATCAAAGCTGT-AAACTACAAATGGAGCCCCAGATGCAGTCCATGA 1677

Qy       7064 C-TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
Db       1678 CTTAAGTTCTACCACAGACCCCTGGACCGGCCTGCTGGCCCATGATCCGATGTTAATGAC 1737

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
Db       1738 ATCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTACGCCCCAATTCA 1797

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
Db       1798 GCAGGAAGCAGTTAGAGCAGTCATCGGCCAACCTCCCCAATAGCACTTAGGTTTTCCTGT 1857

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
Db       1858 TGAGAGGGGGCACTAAGAGACAGGACTAGCTGGATTTCCTAGGACAACTAAGAATCCCTA 1917

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
Db       1918 AGGCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACAGGGGGCTTGCAACTTAGCTCAC 1977

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
Db       1978 ACCCGACCAATCAGAGAGCTCACTAAAATGCTGATTAGGCAAAAACAGAAGGTAAAGAAA 2037

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
Db       2038 TAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATGGGGATATAAACCC 2097

Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
Db       2098 AGGCATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 2157

Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
Db       2158 GTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 2197


RESULT 92
DQ256482
LOCUS       DQ256482               2342 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Cercocebus torquatus atys endogenous virus ERVWE1 syncytin 1
            pseudogene, partial sequence.
ACCESSION   DQ256482
VERSION     DQ256482.1  GI:82407263
KEYWORDS    .
SOURCE      Cercocebus torquatus atys (sooty mangabey)
  ORGANISM  Cercocebus torquatus atys
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Cercocebus.
REFERENCE   1  (bases 1 to 2342)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2342)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2342
                     /organism="Cercocebus torquatus atys"
                     /mol_type="genomic DNA"
                     /sub_species="atys"
                     /db_xref="taxon:9531"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2198
                     /gene="syncytin 1"
                     /pseudo
     LTR             1872. .2298

  Query Match             26.1%;  Score 1957.1;  DB 1;  Length 2342;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches 2042;  Conservative   33;  Mismatches  122;  Indels    3;  Gaps    3;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
Db          1 TTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCATGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
Db         61 AGTCAAACCCTGGATACAGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
Db        121 CTGTGAACCTCTGGAGGATTTGCGCCTGCTCTTCAAACGACAACCAGGAGGAAAGTAACT 180

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
Db        181 AAAATCGTAAAT-CCCTATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTCTTAC 239

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
Db        240 CCCCTTTCACTCTCACTGCACCCCCTCCATGCCTCTGTACGACCAGTAGCTCCCCTTACC 299

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
Db        300 AAGAGTTTCTATGGAGAATGCAGCGCCCTGGAAATATTGATGCTTCATTGTATAGGAGTT 359

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
Db        360 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGTTATAATG 419

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
Db        420 CTGCCACTCTTTGCATGCAATCAAATACTCACTATTGGACAGGAAAAATGATTAATCCTA 479

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
Db        480 GTTGTCCTGGAAGACTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGT 539

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
Db        540 CTGATGGGAGTGGAGTTCAAGATCAGGCAAGAGAGAAACACATAGAAGAAGTAATCTCCC 599

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
Db        600 AACTGACCCGGGTACATAGCACCTCTAGCCCCTGCAGAGGACTAGATCTCTCAAAACTAC 659

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
Db        660 ATGAAACCCTCCGTATCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCCTTGGGC 719

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
Db        720 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCA 779
```

```
Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGCCATACATTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 839

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             |||||||  ||| |||||||||||||||||||||| :|||||||||| ||||||||||||
Db        840 CCACTTCTGTTCTAGTAGGACCTCTTGTTTCCAACCTGGAAATAACTCATACCTCAAACC 899

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||| |||||||  |||||  | ||||||||||||||||||||||||||||||
Db        900 TCACCTGTGTAAATTTTAGCAGTACTATAGACACAACCAACTCCCAATGCATCAGGTGGG 959

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||  ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        960 TAAGTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGCGGTACCT 1019

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAGACTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1079

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             | ||:|||:|||||||||||||||||||||||||||| ||||||||||  ||||||||||
Db       1080 CACCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATACATAAGCCCCGCA 1139

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGGTGGACTAGGTA 1199

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1259

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             ||||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 ATGGTGACATAGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 1319

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             ||||||||||||||||:|||||  |||||||||||||:|  ||||  ||  ||:|||||||
Db       1320 TAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTAGACTTGTTAACTGCCGAAAGAGGGG 1379

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Db       1380 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCA 1439

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             ||||||||||:||||||||  ||||||||||||||||:|||| |:||||||| |||||
Db       1440 GTGAGAAAGTTAAAGAAATTAAAGATCGAATACAACGTAGAGCAAAGGAGCTTCAAAACA 1499

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||| |:|||||||||| |||||||||||||||||||| ||
Db       1500 CTGGACCCTGGGGCCTCCTCAGTCAATGGATGCCCCGGATTCTCCCCTTCTTAGGACTTC 1559

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||||||||||||||||| |||||||||||||:||||| ||||||| |
Db       1560 TAGCAGCTATAATATTGCTACTCCTCTTTGTACCCTGTATCTTTAACCTCTTTGTTAAGT 1619

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063
             ||||||||||  ||||||  |||||||  |||||||||||||||||||  |||||||||| ||
Db       1620 TTGTCTCTTCTAGAATCAAAGCTGT-AAACTACAAATGGAGCCCCAGATGCAGTCCATGA 1678

Qy       7064 C-TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
             | ||||  ||||||||||||||||||||||||||||:|||||||  ||||  ||||||||||
Db       1679 CTTAAGTTCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCATGATCCGATGTTAATGAC 1738

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
             ||||||||||  |||||||||||||||||  ||||||||||  ||||||  ||||||||||
Db       1739 ATCAAAGGCACGCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCA 1798

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
             |||||||||||||||||  ||: |||||||||||||| || ||||||||||||||||||||
Db       1799 GCAGGAAGCAGTTAGAGCAGTCATCGGCCAACCTCCCTAATAGCACTTAGGTTTTCCTGT 1858

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
             |||||  |||| ||| |||||||||||||||||||||||||||||   |:|||||||:||
Db       1859 TGAGAGGGGGCACTAAGAGACAGGACTAGCTGGATTTCCTAGGACAACTAAGAATCCCTA 1918

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
             || ||||:|||||||||||||| |||||||||||||||| |||||||||||||||:||||
Db       1919 AGGCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACAGGGGGCTTGCAACTTAGCTCAC 1978

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
             |||  |||||||||||||||||||||||||||||||| ||||||||| |||| ||||||||||
Db       1979 ACCCGACCAATCAGAGAGCTCACTAAAATGCTGATTAGGCAAAAACAGAAGGTAAAGAAA 2038

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
             ||||||||||:|||||:|||||||||||||| ||||||||| ||||||||||||||||||
Db       2039 TAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCC 2098

Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
             | |  ||| ||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db       2099 AGGCATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 2158

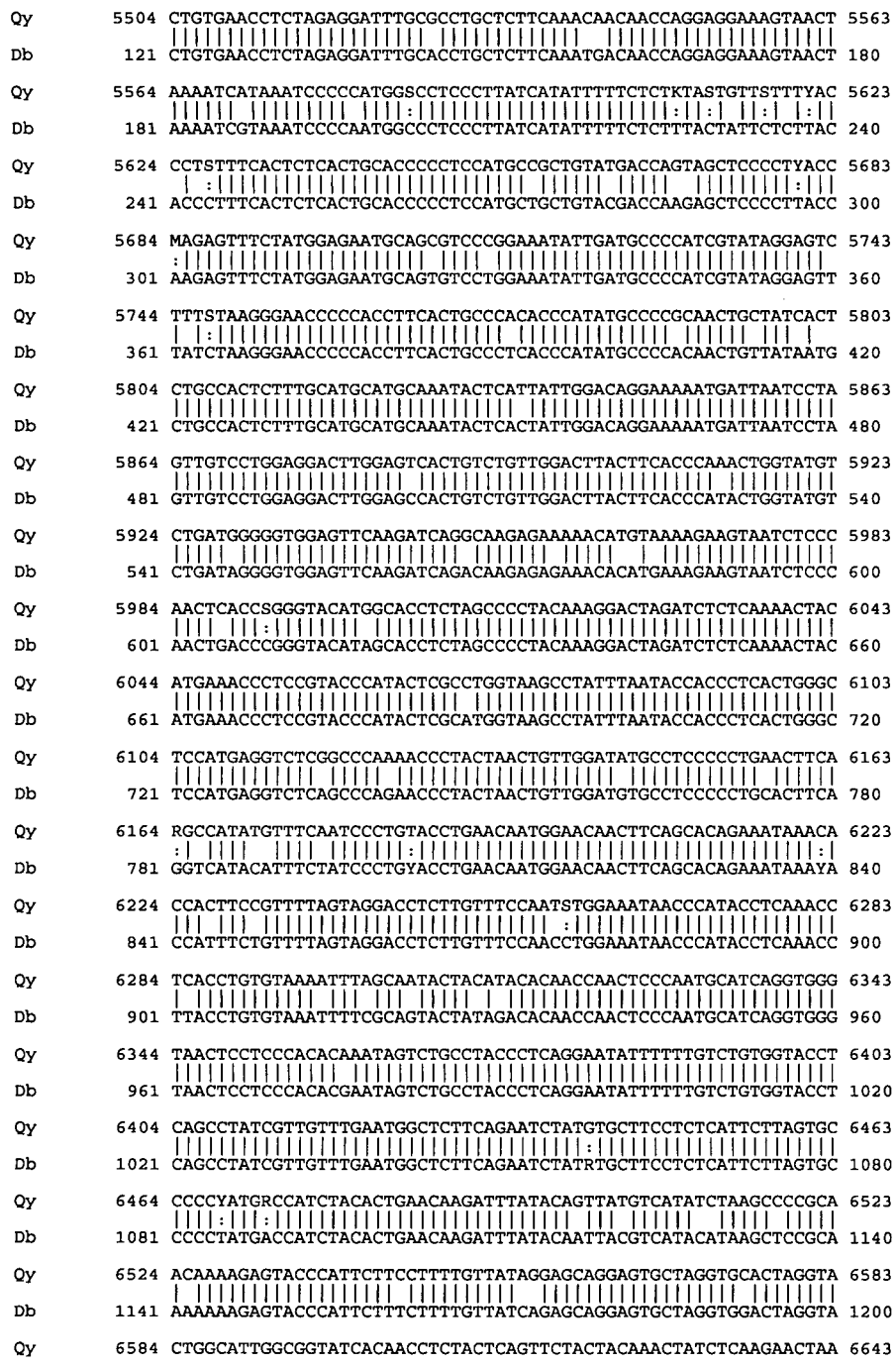
Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||  |||||||||||||||||||||||:||||:
Db       2159 GTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 2198
```

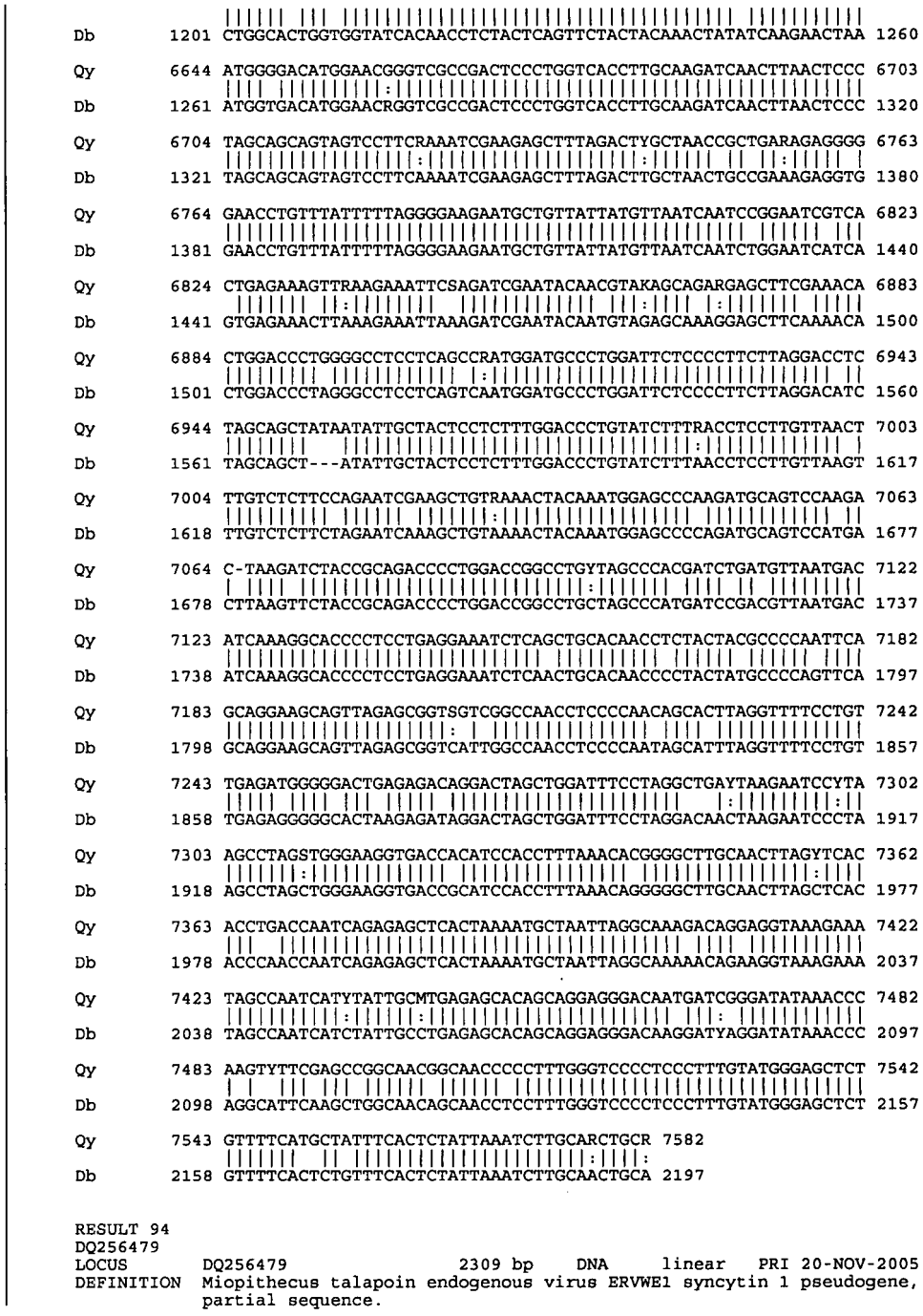
```
RESULT 93
DQ256477
LOCUS       DQ256477               2344 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Colobus guereza kikuyuensis endogenous virus ERVWE1 syncytin 1
            pseudogene, complete sequence.
ACCESSION   DQ256477
VERSION     DQ256477.1  GI:82407258
KEYWORDS    .
SOURCE      Colobus guereza kikuyuensis
  ORGANISM  Colobus guereza kikuyuensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Colobinae; Colobus.
REFERENCE   1  (bases 1 to 2344)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2344)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2344
                     /organism="Colobus guereza kikuyuensis"
                     /mol_type="genomic DNA"
                     /sub_species="kikuyuensis"
                     /db_xref="taxon:132548"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2197
                     /gene="syncytin 1"
                     /pseudo
     LTR             1871. .2300

  Query Match             26.0%;  Score 1950.4;  DB 1;  Length 2344;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches 2027;  Conservative   39;  Mismatches  130;  Indels    4;  Gaps    2;
```

```
Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
             |||||:||| ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          1 TTTATCTACTCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
             ||||||| |||| |||| ||||||||||||||||||||||||||||||||||||||||| ||
Db         61 AGTCAAACCCTGTATACAGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTACTC 120
```





RESULT 94
DQ256479
LOCUS DQ256479 2309 bp DNA linear PRI 20-NOV-2005
DEFINITION Miopithecus talapoin endogenous virus ERVWE1 syncytin 1 pseudogene, partial sequence.

```
ACCESSION   DQ256479
VERSION     DQ256479.1  GI:82407260
KEYWORDS    .
SOURCE      Miopithecus talapoin (talapoin)
  ORGANISM  Miopithecus talapoin
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Miopithecus.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Miopithecus talapoin"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:36231"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2162
                     /gene="syncytin 1"
                     /pseudo
     LTR             1871. .2265

  Query Match             25.5%;  Score 1906.3;  DB 1;  Length 2309;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches 2015;  Conservative   32;  Mismatches  114;  Indels   39;  Gaps    4;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
             |||||:|||||||||||||||||||||||| |||||||||||| |||||||||||||||| |
Db          1 TTTATCTACCCCAACTGCGGTTAAAGTGGGTGGAGTGGAGTCATGGATACATCACACTCG 60

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
             ||||||| |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTCAAACCCTGGATACAGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 120

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
             |||||||||||||||||||||| ||||||||||||  ||||||||||||||||||||||
Db        121 CTGTGAACCTCTAGAGGATTTGCACCTGCTCTTCAAATGACAACCAGGAGGAAAGTAACT 180

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
             |||||||||||| ||||||||:|||||||||||||||||||| |||:||:||||:| |:||
Db        181 AAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTTTCTTTACTGTTCTCTTAC 239

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
             || :|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        240 CCCCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTACGACCAGTAGCTCCCCTTACC 299

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
             :|||||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db        300 AAGAGTTTCTATGGAGAATGCAGCGTCCTGGAAATATTGATGCCCCATTGTATAGGAGTT 359

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        360 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGTTATAATG 419

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             ||||||||||||||||||  |||||||||| ||||||||||||||||||||||||||||||
Db        420 CTGCCACTCTTTGCATGCAATCAAATACTCACTATTGGACAGGAAAAATGATTAATCCTA 479

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||| |||||||||||||||||||||||||||| ||  ||||||
Db        480 GTTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGT 539

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||||||| |||||  ||||||||||||||||||

Db        540 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACACATAAAAGAAGTAATCTCCC 599

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| | |:||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        600 AACTAAACCGGGTACATAGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 659

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        660 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCATTGGGC 719

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        720 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCA 779

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||  |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        780 GGCCATACATTTCAATCCCTGTACCTGAACAAAGGAACAACTTCAGCACAGAAATAAACG 839

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||  |||||||||||||||||||||||||| :|||||||||||| ||| |||||||
Db        840 ACACTTCTGTTTTAGTAGGACCTCTTGTTTCCAACCTGGAAATAACCCGTACTTCAAACC 899

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||| ||||||| ||||| | | ||||||||||||||||||||||||||||||
Db        900 TCACCTGTGTAAATTTTAGCAGTACTATAGATACAACCAACTCCCAATGCATCAGGTGGG 959

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        960 TAACTCCTCCCACACGAATAGTCTGCCTATCCTCAGGAATATTTTTTGTCTGTGGTACCT 1019

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 1079

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             | ||:|||:|||||||||||||||||||||||||||||| |||||||||| |||||||||
Db       1080 CACCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATATAAGCCCCGCA 1139

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db       1140 ACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGGTGGACTAGGTA 1199

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 1259

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||||||||||||| |  ||||||||||||||||||||||||||||||||
Db       1260 ATGGTGACATGGAACGGGTCGCCAA--CCCTGGTCACCTTGCAAGATCAACTTAACTCCC 1317

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:|||||| ||  |:|||||||
Db       1318 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCCAAAAGAGGGG 1377

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             ||| |||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Db       1378 GAATCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGAATCATCA 1437

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             ||||||||||:|||||||  ||||||||||||| |||:|||| |:||||||| |||||
Db       1438 GTGAGAAAGTTAAAGAAATTAAAGATCGAATACAATGTAGAGCAAAGGAGCTTCAAAACA 1497

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||||:|||||||||  ||||||||||||||||||||| ||
Db       1498 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCACGGATTCTCCCCTTCTTAGGACTTC 1557

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||||||| ||||||||||||||||||||:|||||||||||||| |
Db       1558 TAGCAGCTATAATATTGCTACTTCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1617

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGA 7063
             |||||||||| |||||| |||||||:||||||||||||||||| |||||||||||| ||
Db       1618 TTGTCTCTTCTAGAATCAAAGCTGTAAAACTACAAATGGAGCCCCAGATGCAGTCCATGA 1677
```

```
Qy       7064 C-TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGAC 7122
              | |||| ||||||||||||||||||||||||||||:||||||| |||| ||||||||||||
Db       1678 CTTAAGTTCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCATGATCCGATGTTAATGAC 1737

Qy       7123 ATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCA 7182
              |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db       1738 ATCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTACGCCCCAATTCA 1797

Qy       7183 GCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGT 7242
              |||||||||||||||||  ||: |||||||||||||||||  ||||||||||||||||||||
Db       1798 GCAGGAAGCAGTTAGAGCAGTCATCGGCCAACCTCCCCAATAGCACTTAGGTTTTCCTGT 1857

Qy       7243 TGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTA 7302
              ||||| |||| ||| ||||||||||||||||||||||||||||   |:|| ||||||: |
Db       1858 TGAGAGGGGGCACTAAGAGACAGGACTAGCTGGATTTCCTAGGACAACTAGGAATCCCCA 1917

Qy       7303 AGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAC 7362
              || ||||:|||||||||||||| ||||||||||||||| |||||||||| ||||:||||
Db       1918 AGGCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACAGGGGGCTTGCAAGTTAGCTCAC 1977

Qy       7363 ACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAA 7422
              ||| ||||||||||||||||||||||||||||| |||||||||| |||| ||||||||||
Db       1978 ACCCGACCAATCAGAGAGCTCACTAAAATGCTCATTAGGCAAAAACAGAAGGTAAAGAAA 2037

Qy       7423 TAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCC 7482
              |||||||||                                   ||||||||||||||||
Db       2038 TAGCCAATC-----------------------------------ATCGGGATATAAACCC 2062

Qy       7483 AAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 7542
              |    ||| ||| |||||| ||||||||||||||||||||||||||||||||||||||||
Db       2063 AGACATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCT 2122

Qy       7543 GTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||  |||||||||||||||||||| ||:||||:
Db       2123 GTTTTCACTCTATTTCACTCTATTAAATCTTACAACTGCA 2162
```

```
RESULT 95
DQ256478
LOCUS       DQ256478               2319 bp    DNA     linear   PRI 20-NOV-2005
DEFINITION  Presbytis francoisi endogenous virus ERVWE1 syncytin 1 pseudogene,
            partial sequence.
ACCESSION   DQ256478
VERSION     DQ256478.1  GI:82407259
KEYWORDS    .
SOURCE      Presbytis francoisi
  ORGANISM  Presbytis francoisi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Colobinae; Presbytis.
REFERENCE   1  (bases 1 to 2319)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     The gene of retroviral origin syncytin 1 is specific to hominoids
            and was inactivated in Old World monkeys
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2319)
  AUTHORS   Caceres,M. and Thomas,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2005) Department of Human Genetics, Emory
            University School of Medicine, 615 Michael Street, Suite 301,
            Atlanta, GA 30322, USA
FEATURES             Location/Qualifiers
     source          1. .2319
                     /organism="Presbytis francoisi"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:33549"
                     /note="endogenous retroviral family W, env(C7), member 1
                     endogenous_virus: ERVWE1"
     gene            <1. .2172
                     /gene="syncytin 1"
                     /pseudo
     LTR             1846. .2275
```

```
  Query Match             24.8%;  Score 1859.7;  DB 1;  Length 2319;
  Best Local Similarity   90.6%;  Pred. No. 0;
  Matches 1999;  Conservative   34;  Mismatches  132;  Indels   41;  Gaps    6;

Qy       5384 TTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTG 5443
              |||||:||| |||||||||||||||||     |||||||||||||||||||||| |||||||||
Db          1 TTTATCTACTCCAACTGCGGTTAAAG----TGGAGTGGAGTCTTGGATACGTCACACTTG 56

Qy       5444 AGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 5503
              ||||||| ||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db         57 AGTCAAACCCTGGATATAGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTC 116

Qy       5504 CTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACT 5563
              |||||||||||||||||| |||| |||||||||||  ||||||||||||||||||||||||
Db        117 CTGTGAACCTCTAGAGGATGTGCGTCTGCTCTTCAAATGACAACCAGGAGGAAAGTAACT 176

Qy       5564 AAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYAC 5623
              |||||| ||||| | |||||||:||||||||||||| |||||||||: |:||||:| |:||
Db        177 AAAATCGTAAAT-CTCCATGGCCCTCCCTTATCATGTTTTTCTCTTGACTGTTCTCTTAC 235

Qy       5624 CCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACC 5683
              | :|||||||||||||||||||||||||||||| |||||||||  |||||| |||||||||:|||
Db        236 ACCCTTTCACTCTCACTGCACCCCCTCCAGGCCGCTGTACGACCAGGAGCTCCCCTTACC 295

Qy       5684 MAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTC 5743
              :|||||||||||||||||||| |||| ||||||||||||||||||| |||||||||
Db        296 AAGAGTTTCTATGGAGAATGCAGTGTCCTGGAAATATTGATGCCCCATCATATAGGAGTT 355

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
              | |:|||||||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        356 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGTTATAATG 415

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAA-------AATGATT 5856
              |||||||||||||||||||||||||||||||| |||||||||||||        |||||||
Db        416 CTGCCACTCTTTGCATGCATGCAAATACTCACTATTGGACAGGAAAAATGATTAATGATT 475

Qy       5857 AATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACT 5916
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||
Db        476 AATCCTAGTTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACC 535

Qy       5917 GGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTA 5976
              |||||||||||  |||||||||||||||||||||||||||||||||  |||||  ||||||||||
Db        536 GGTATGTCTGATAGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACACATAAAAGAAGTA 595

Qy       5977 ATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCA 6036
              |||||||||||  |||:|||||||  ||||||||||||||||||||||||||||||||||||
Db        596 ATCTCCCAACTGACCCGGGTACAGAGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCA 655

Qy       6037 AAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTC 6096
              |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||
Db        656 AAACTACATGAAACCCTCCGTACCCACACTCGCGTGGTAAGCCTATTTAATACCACCCTC 715

Qy       6097 ACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTG 6156
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||
Db        716 ACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTAGATGTGCCTCCCCCTG 775

Qy       6157 AACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAA 6216
              ||||||:| ||||  |||| ||||||||||||||||||||||||||||||||||||||||||||
Db        776 CACTTCAGGTCATACATTTCTATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAA 835

Qy       6217 ATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACC 6276
              |||||||||||||| |||||||||||||||||||||||||||| :| ||||||||||||||||
Db        836 ATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAACCTCGAAATAACCCATACC 895

Qy       6277 TCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATC 6336
              |||||||||||||||||||| ||| ||| ||||| | ||||||||||||||||||||||||
Db        896 TCAAACCTCACCTGTGTAAATTTTCGCAGTACTATAGACACAACCAACTCCCAATGCATC 955

Qy       6337 AGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGT 6396
              | |||||||||||||||||| |||||||||||| |||||||||||||||||||||||| |||
Db        956 AAGTGGGTAACTCCTCCCACACGAATAGTCTGCCTAACCTCAGGAATATTTTTTGTTTGT 1015
```

```
Qy       6397 GGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTC 6456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1016 GGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTC 1075

Qy       6457 TTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAG 6516
             ||||||||||:|||:|||||||||||||||||||||||||||| ||| ||||||| ||||
Db       1076 TTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTACGTCATATGTAAG 1135

Qy       6517 CCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCA 6576
             |||||||||||||||||||||||||||||||||||||  |||||||||||||||||| |
Db       1136 CCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGTGGA 1195

Qy       6577 CTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAA 6636
             |||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||
Db       1196 CTAGGTACTGGCACTGGTGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAA 1255

Qy       6637 GAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTT 6696
             |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1256 GAACTAAATGGTGACATGGAACAGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTT 1315

Qy       6697 AACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGAR 6756
             ||| ||||||||||||||||||:||||||||||||||||||:|||||| || ||:
Db       1316 AAC-CCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACTGCCGAA 1374

Qy       6757 AGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGA 6816
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1375 AGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCTGGA 1434

Qy       6817 ATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTT 6876
             |||  || ||||||| ||:|||| |||  ||||| ||||||||||:|||| |:||||||
Db       1435 ATCACCAGTGAGAAACTTAAAGACATTAAAGATCAAATACAACGTAGAGCAAAGGAGCTT 1494

Qy       6877 CGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTA 6936
             | ||| |||||||||||  |||||| |||| |:||||||||||||||||||| ||||||||
Db       1495 CAAAATACTGGACCCTGCGGCCTCTTCAGTCAATGGATGCCCTGGATTCTCTCCTTCTTA 1554

Qy       6937 GGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTT 6996
             |||||||||||||||   ||||||||||||||||||||||||||||||:|||||||||
Db       1555 GGACCTCTAGCAGCT---ATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTT 1611

Qy       6997 GTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAG 7056
             ||||| |||||||||||  |||||| |||||||:|||||| |
Db       1612 GTTAAGTTTGTCTCTTCTAGAATCAAAGCTGTAAAACTAGA------------------- 1652

Qy       7057 TCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTT 7116
                  |||| |||||||||||||||||||||||||||||:||||||| |||| ||| ||
Db       1653 ------CTTAAGTTCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCATGATCCGATATT 1706

Qy       7117 AATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCC 7176
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||
Db       1707 AATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACTATGCCCC 1766

Qy       7177 AATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTT 7236
             |||||||||||||||||| ||||||:   |||||||||||||| ||||||||||||
Db       1767 AATTCAGCAGGAAGCAGTTACAGCGGTCACTGGCCAACCTCCCCAATAGCACTTAGGTTT 1826

Qy       7237 TCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAA 7296
             |||||||||| |||| ||| ||||| ||||||||| ||||||||||   |:||||||
Db       1827 TCCTGTTGAGAGGGGGCACTAAGAGATAGGACTAGCCGGATTTCCTAGGACAACTAAGAA 1886

Qy       7297 TCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTA 7356
             |||:|||||||:|||||||||||| |||||||||||||||| ||||||||||||||||||
Db       1887 TCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACAGGGGGCTTGCAACTTA 1946

Qy       7357 GYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTA 7416
             |:|||||||  ||||||||||||||||||||||||||||||||||||| |||| |||||
Db       1947 GCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGAAGGTA 2006

Qy       7417 AAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATAT 7476
             |||||||||||| ||||:||||||:|||||||||||||||||||||||| ||| |||||||
Db       2007 AAGAAATAGCCACTCATCTATTGCCTGAGAGCACAGCAGGAGGGACAAGGATTGGGATAT 2066

Qy       7477 AAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGG 7536
             ||||||| |  ||| ||| |||||| |||||| |||||||||||||||||||||||||||
Db       2067 AAACCCAGGCATTCAAGCTGGCAACAGCAACCTCCTTTGGGTCCCCTCCCTTTGTATGGG 2126

Qy       7537 AGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||  |||||||||||||||||||||||||||||:||||:
Db       2127 AGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTGCA 2172
```

RESULT 96
AF135487

```
LOCUS       AF135487               2309 bp    DNA     linear   PRI 27-OCT-1999
DEFINITION  Homo sapiens retroviral-related sequence identified from
            schizophrenia by RDA, mapped to multiple sites.
ACCESSION   AF135487
VERSION     AF135487.1  GI:4928644
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Deb-Rinker,P., Klempan,T.A., O'Reilly,R.L., Torrey,E.F. and
            Singh,S.M.
  TITLE     Molecular characterization of a MSRV-like sequence identified by
            RDA from monozygotic twin pairs discordant for schizophrenia
  JOURNAL   Genomics 61 (2), 133-144 (1999)
   PUBMED   10534399
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Deb-Rinker,P., Klempan,T.A., O'Reilly,R.L. and Singh,S.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-1999) Molecular Genetics, Univ. Of Western
            Ontario, 307 Western Science Centre, London, ON N6A 5B7, Canada
COMMENT     A 127 bp retroviral-related region of this sequence was identified
            by amplicon-based RDA from 3 pairs of monozygotic twins discordant
            for schizophrenia, where the amplicon from the affected twin was
            used as the tester. The sequence was expanded to 2.3 kb using Human
            genome walker PCR (Clontech) and shows similarity to a multiple
            sclerosis associated retrovirus polyprotein mRNA (AF009668). A
            region of this sequence was used as a probe for screening a BAC
            library, which identified 26  positive BAC clones, representing 8
            different types. 6 of these were used for FISH mapping. The results
            argue for multiple and dispersed genomic sites.
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1, 2, 3, 6, 12, 14"
                     /map="1p32.2-p34.2, 2q11.2, 3p12-p13, 6q15-q16.2,
                     12q13.2-q13.3, 14q21"
     repeat_region   <1. .>2309
                     /note="retroviral-related sequence identified from
                     schizophrenia"
```

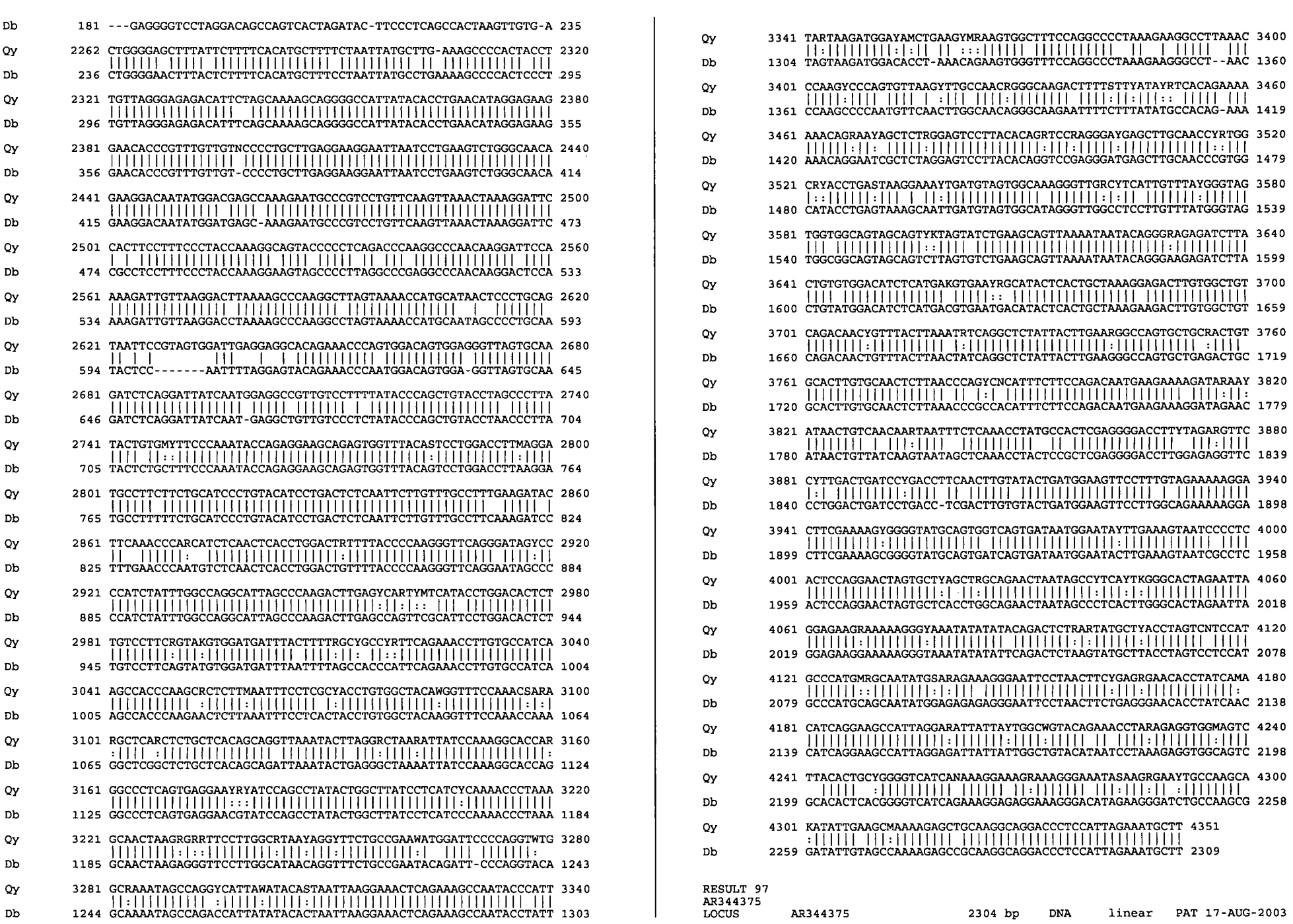
```
  Query Match             24.8%;  Score 1859.6;  DB 1;  Length 2309;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 2037;  Conservative  120;  Mismatches  148;  Indels   26;  Gaps   15;

Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             |||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGCGCCAGCCCATGCCATCACCCTCACAG 60

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| |||||||||||||||||||||||||||||| | |   |||||||||||||||| |
Db         61 AGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCCTGGACACTGGCG 120

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             ||| ||||| |||||||||| ||| |||||||||||||||||||||||||||||||| |
Db        121 CGGCCTTCTCAGTCTTACTCTCCTGACCCGGACAACTGTCCTCCAGATCTGTCACTATCC 180

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                 | |||  | || ||| | |||||||||||||| ||: | |||||||||||| || |
```



RESULT 97
AR344375
LOCUS AR344375 2304 bp DNA linear PAT 17-AUG-2003

```
DEFINITION  Sequence 87 from patent US 6582703.
ACCESSION   AR344375
VERSION     AR344375.1  GI:33740316
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2304)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 87 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .2304
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             24.7%;  Score 1851.9;  DB 1;  Length 2304;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches 2032;  Conservative  119;  Mismatches  150;  Indels   29;  Gaps   15;

Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             ||||||||||||||||||||||||| |||||||| |||| ||||||||||||||||  |||
Db             1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| ||||||| |||||||||||| ||||||||| |   |||||||||||||||| |
Db            59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             | | |||||| ||||||||| | |||||||| ||||| |||||||||||||||||||||| |
Db            119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                | |||  | || ||| | ||||||||| |||| ||:|| |||||||||||||| || |
Db            179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
             ||||||| ||||| |||||||||||||||||||||||||| |||||||||||||| ||||
Db            234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db            294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
             || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||||| ||
Db            354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
             |||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||
Db            413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
              | |||||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db            472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
             |||||||||||||| |||||||||||||||| ||||||||||||||||||  |  ||||||| |
Db            532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
             | | |      |||   | | |  |||||||  |||||||||| ||||||||||||
Db            592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
             |||||||||||||| ||| ||||| ||| | | | |||||||||||||| ||||||||||
Db            644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
```

```
                 ||| ||::||||| |||||||||||||||||||||||||||:|||||||||||:|||||
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861
             ||||| |||||||||||||||| |||||||||||||||||||||||||||||||||| ||
Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921
             |  ||||||:| |||||||||||||||||:|||||||||||||||||||||||||:|||
Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981
             ||||||||||||||||||||||||||||||||||||:||:|::|||||||||||||||||
Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041
             |||||||:||| |||||||||||||||||:| :|||::|||||||||||||||||||||
Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101
             |||||||||| :|||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:
Db        1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161
             |||| :|||||||||||| || ||| |||||||||:||||:|||||||||||||||||:|
Db        1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221
             |||||||||||||||:::|||||||||||||||||||||||||||||||:|||||||||||
Db        1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281
             ||||||||:|::|||||||||:|||:|||:|||||||||:|  |||| |||||||:
Db        1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341
             |  |||||||||  :|||||:|||||:|||||| ||||||||||||||||||||||| ||||
Db        1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401
             |:|||||||||:|:||    :::|||||||||||||||||  ||  | | |||  ||||
Db        1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461
             ||||:|||||||||  ||:||||||||:|||||||  ||||:||:|||::  |||||  ||||
Db        1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521
             |||||:||:||||||:|||||||||||| |||:||  |||||:|||||||||||::|||
Db        1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581
             ::||||||:||||||||:|||||||||||||||||||||:|:||||||||||:||||| |
Db        1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641
             || |||||||||||::|||||||||||||||||||||||||||||||||:|||||||||||
Db        1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701
             ||||||||||||||||||:||||||::|||||||||||||||||||||||||||||| ||||
Db        1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761
             |||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|||| |
Db        1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821
             |||||||||||||||||| || |:| |||||||||||||||||||||||||||||||:||:|
Db        1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881
             ||||||||||||:||||| ||||||||||||  |||||||||||||||:||||:|||||
```

```
Db       1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941
             :|||||||||||:|||| |||||||||||||||||||||||||| | ||||||||||
Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001
             |||||||||:|||||||||||| |||||||||||||||||:|||||||||||| |||||
Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061
             |||||||||||||||||:| ||:||||||||||||||:|||:|:|||||||||||||||
Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121
             ||||||:||||||||:|||||||||| ||||||||:|:||||||:|||||||| ||||||
Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181
             ||||||::||||||||:|:||| |||||||||||||||||:|||:||||||||||||: |
Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241
             |||||||||||||||||:|||||||:||||:|||||||||||||:||||||||:|||||
Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301
             ||||||||: |||||||||  ||| | ||:||||||||||||:|||: ||: ||||||| :
Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
             ||||||||||:|||| ||| |||||||||||| ||||||||||||||||
Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 98
AF009668
LOCUS       AF009668               2304 bp    mRNA    linear   VRL 12-AUG-1997
DEFINITION  Multiple sclerosis associated retrovirus polyprotein (pol) mRNA,
            partial cds.
ACCESSION   AF009668
VERSION     AF009668.1  GI:2323286
KEYWORDS    .
SOURCE      Multiple sclerosis associated retrovirus
  ORGANISM  Multiple sclerosis associated retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 2304)
  AUTHORS   Perron,H., Garson,J.A., Bedin,F., Beseme,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Mallet,F., Tuke,P.W., Voisset,C., Blond,J.L.,
            Lalande,B., Seigneurin,J.M., Mandrand,B. and The collaborative
            research group on MS.
  TITLE     Molecular identification of a novel retrovirus repeatedly isolated
            from patients with multiple sclerosis. The Collaborative Research
            Group on Multiple Sclerosis
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 94 (14), 7583-7588 (1997)
   PUBMED   9207135
REFERENCE   2  (bases 1 to 2304)
  AUTHORS   Perron,H., Bedin,F., Beseme,F., Mallet,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Garson,J.A. and Tuke,P.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1997) UMR103 / E.N.S., BIOMERIEUX, 46, allee
            d'Italie, Lyon CEDEX 07 69364, France
COMMENT     International patent WO-97/06260 and other patents pending.
FEATURES             Location/Qualifiers
     source          1. .2304
                     /organism="Multiple sclerosis associated retrovirus"
                     /mol_type="mRNA"
                     /db_xref="taxon:62805"
     gene            <1. .>2304
                     /gene="pol"
     CDS             <1. .>2304
                     /gene="pol"
                     /note="contains protease and reverse transcriptase"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="AAB66528.1"
                     /db_xref="GI:2323287"
                     /translation="SSSRTEGARGKCQPMPSPSEPRVCLTIESQEVNCLLDTGAAFSV
                     LLSCPRQLSSRSVTIRGVLGQPVTTYFSQPLSCDWGTLLFSHAFLIMPESPTPLLGRD
                     ILAKAGAIIHLNIGKGIPICCPLLEEGINPEVWAIEGQYGQAKNARPVQVKLKDSASF
                     PYQRKYPLRPEALQGXQKIVKDLKAQGLVKPCSSPCNTPILGVRKPNGQWRLVQDLRI
                     INEAVFPLYPAVSSPYTLLSLIPEEAEWFTVLDLKDAFFCIPVRPDSQFLFAFEDPLN
                     PTSQLTWTVLPQGFRDSPHLFGQALAQDLSQFSYLDTLVLQYVDDLLLVARSETLCHQ
                     ATQELLTFLTTCGYKVSKPKARLCSQEIRYLGLKLSKGTRALSEERIQPILAYPHPKT
                     LKQLRGFLGITGFCRKQIPRYTPIARPLYTLIRETQKANTYLVRWTPTEVAFQALKKA
                     LTQAPVFSLPTGQDFSLYATEKTGIALGVLTQVSGMSLQPVVYLSKEIDVVAKGWPHC
                     LWVMAAVAVLVSEAVKIIQGRDLTVWTSHDVNGILTAKGDLWLSDNHLLNYQALLLEE
                     PVLRLRTCATLKPATFLPDNEEKIEHNCQQVIAQTYAARGDLLEVPLTDPDLNLYTDG
                     SSLAEKGLRKAGYAVISDNGILESNRLTPGTSAHLAELIALTWALELGEGKRVNIYSD
                     SKYAYLVLHAHAAIWREREFLTSEGTPINHQEAIRRLLLAVQKPKEVAVLHCQGHQEE
                     EEREIEGNRQADIEAKKAARQDSPLEML"

  Query Match             24.7%;  Score 1851.9;  DB 1;  Length 2304;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches 2032;  Conservative  119;  Mismatches  150;  Indels   29;  Gaps   15;

Qy       2025 TCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAG 2084
             |||||||||||||||||||||||| |||||||| |||| |||||||||||||||  |||
Db          1 TCCAGCAGCAGGACTGAGGGTGCCCGGGGCAAGTGCCAGCCCATGCCATCACCCT--CAG 58

Qy       2085 AGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTG 2141
             ||||| ||||||| |||||||||||| |||||||| |   ||||||||||||||||| |
Db          59 AGCCCCGGGTATGTTTGACCATTGAGAGCCAGGAAGTTAACTGTCTCCTGGACACTGGCG 118

Qy       2142 CGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTC 2201
             | | ||||| |||||||||  | ||||||| ||||| ||||||||||||||||||||||| |
Db         119 CAGCCTTCTCAGTCTTACTTTCCTGTCCCAGACAATTGTCCTCCAGATCTGTCACTATCC 178

Qy       2202 TGAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAA 2261
                | |||  | || ||| | |||||||||| |||| ||:|| |||||||||||||| || |
Db         179 ---GAGGGGTCCTAGGACAGCCAGTCACTACATAC-TTCTCTCAGCCACTAAGTTGTG-A 233

Qy       2262 CTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTT 2321
             ||||||| ||||| |||||||||||||||||||||||||| |||||||||||||| ||||
Db         234 CTGGGGAACTTTACTCTTTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTT 293

Qy       2322 GTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGG 2381
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db         294 GTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAAAAGG 353

Qy       2382 AACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAG 2441
             || |||| |||| ||| |||||||||||||||||||||||||||||||||||||||||| ||
Db         354 AATACCCATTTGCTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAATAG 412

Qy       2442 AAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCC 2501
             |||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db         413 AAGGACAATATGGACAAG-CAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCT 471

Qy       2502 ACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAA 2561
               | |||||||||||||||| ||||||| || |||||| |||||| |||||||   |||
Db         472 GCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTAGACCCGAGGCCCTACAAGGANCTCAA 531

Qy       2562 AAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGT 2621
             |||||||||||||| |||||||||||||| |||||||||||||| |  ||||||| |
Db         532 AAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGCCCCTGCAAT 591

Qy       2622 AATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAG 2681
             | | |        |||   | | | ||||||||  |||||||||| ||||||||||||
Db         592 ACTCC-------AATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAG 643

Qy       2682 ATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTAT 2741
             |||||||||||||| ||| ||||| ||| | | | |||||||||||||| ||||||||||
Db         644 ATCTCAGGATTATTAAT-GAGGCTGTTTTTCCTCTATACCCAGCTGTATCTAGCCCTTAT 702

Qy       2742 ACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGAT 2801
             ||| ||::||||| |||||||||||||||||||||||||||:||||||||||:|||||
```

```
Db        703 ACTCTGCTTTCCCTAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGAT 762

Qy       2802 GCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACT 2861

Db        763 GCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCT 822

Qy       2862 TCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCC 2921

Db        823 TTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCC 882

Qy       2922 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTT 2981

Db        883 CATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTT 942

Qy       2982 GTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAA 3041

Db        943 GTCCTTCAGTACGTGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAA 1002

Qy       3042 GCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARAR 3101

Db       1003 GCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAG 1062

Qy       3102 GCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARG 3161

Db       1063 GCTCGGCTCTGCTCACAGGAGATTAGATACTTAGGGCTAAAATTATCCAAAGGCACCAGG 1122

Qy       3162 GCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAG 3221

Db       1123 GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAG 1182

Qy       3222 CAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGG 3281

Db       1183 CAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACAC 1241

Qy       3282 CRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTT 3341

Db       1242 CCCAATAGCCAGACCATTATATACACTAATTAGGGAAACTCAGAAAGCCAATACCTATTT 1301

Qy       3342 ARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACC 3401

Db       1302 AGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACC 1356

Qy       3402 CAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAA 3461

Db       1357 CAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAA 1415

Qy       3462 AACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGC 3521

Db       1416 AACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT 1475

Qy       3522 RYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGT 3581

Db       1476 ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAAT 1535

Qy       3582 GGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTAC 3641

Db       1536 GGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTAC 1595

Qy       3642 TGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTC 3701

Db       1596 TGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTC 1655

Qy       3702 AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTG 3761

Db       1656 AGACAACCATTTACTTAATTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGAGACTGCG 1715

Qy       3762 CACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYA 3821

Db       1716 CACTTGTGCAACTCTTAAACCCGCCACATTTCTTCCAGACAATGAAGAAAAGATAGAACA 1775

Qy       3822 TAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCC 3881

Db       1776 TAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCGAGGGGACCTTCTAGAGGTTCC 1835

Qy       3882 YTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGAC 3941

Db       1836 CTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTGGCAGAAAAAGGAC 1894

Qy       3942 TTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCA 4001

Db       1895 TTCGAAAAGCGGGGTATGCAGTGATCAGTGATAATGGAATACTTGAAAGTAATCGCCTCA 1954

Qy       4002 CTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAG 4061

Db       1955 CTCCAGGAACTAGTGCTCACCTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAG 2014

Qy       4062 GAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATG 4121

Db       2015 GAGAAGGAAAAAGGGTAAATATATATTCAGACTCTAAGTATGCTTACCTAGTCCTCCATG 2074

Qy       4122 CCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMAC 4181

Db       2075 CCCATGCAGCAATATGGAGAGAGAGGGAATTCCTAACTTCTGAGGGAACACCTATCAACC 2134

Qy       4182 ATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCT 4241

Db       2135 ATCAGGAAGCCATTAGGAGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCT 2194

Qy       4242 TACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAK 4301

Db       2195 TACACTGCCAGGGTCATCAGGAAGAAGAGGAAAGGGAAATAGAAGGCAATCGCCAAGCGG 2254

Qy       4302 ATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351

Db       2255 ATATTGAAGCAAAAAAAGCCGCAAGGCAGGACTCTCCATTAGAAATGCTT 2304


RESULT 99
BD196238
LOCUS       BD196238               2006 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196238
VERSION     BD196238.1  GI:33006008
KEYWORDS    JP 2002512530-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 4 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/4
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
           MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
           Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC           diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .2006
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
```

Query Match 24.7%; Score 1848.6; DB 1; Length 2006;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1949; Conservative 56; Mismatches 1; Indels 206; Gaps 3;

```
Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
             ||:||||:|||||||||||||||:|||:|:|||||||||||||||||||:|||||||||
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
             :|||||||||||||||||||:|:|||||:||||||||  ||||||||||||::|||||||
Db        121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
             |:|:|||||||||||||||||||||:|||:||||||||||||:||||||||||||||||||
Db        181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
             |||:|||||||:||||:|||||||||||||:||||||||:||||||||||||||:||||||
Db        241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
             ||||||||||||||:|
Db        301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376

Db        317 ------------------------------------------------------------ 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436

Db        317 ------------------------------------------------------------ 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
                                                 ||||||||||||||||||||||
Db        317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
             |:|||:|||||||||||||:|||||||:|:|||||||||||||||:|||||||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:|||||:||||||||||||||||||:|:||:|||||||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db         818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
             ||||||||||||||||||||||||||::||||||||||||||||||||||||||:||||
Db        1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||||||||||||||||:||:||||:|||:||||:|||||||||
Db        1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             ||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||
Db        1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
             |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
```

```
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006


RESULT 100
BD267480
LOCUS       BD267480               2006 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267480
VERSION     BD267480.1  GI:33077248
KEYWORDS    JP 2002534980-A/21.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 21 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/21
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .2006
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             24.7%;  Score 1848.6;  DB 1;  Length 2006;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1949;  Conservative   56;  Mismatches    1;  Indels  206;  Gaps    3;

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db        121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db        181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db        241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316

Db        301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db        317 ---------------------------------------------------------- 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db        317 ---------------------------------------------------------- 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
Db        317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
Db       1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
Db       1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
```

```
Db       1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db       1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||||||||||||||:||:||||:|||:||||||:|||||||||
Db       1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             ||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||
Db       1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006


RESULT 101
AX000959
LOCUS       AX000959               2006 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 4 from Patent WO9902696.
ACCESSION   AX000959
VERSION     AX000959.1  GI:7241201
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 4 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             24.7%;  Score 1848.6;  DB 1;  Length 2006;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1949;  Conservative   56;  Mismatches    1;  Indels  206;  Gaps    3;

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
             ||:||||:|||||||||||||||:|||:|:|||||||||||||||||:|||||||||||||
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
              :|||||||||||||||||:|:||||||:|||||||| ||||||||||||::|||||||
Db         121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
              |:|:|||||||||||||||||||:|||:||||||||||||:|||||||||||||||||||
Db         181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
              |||:|||||||:||||:|||||||||||||:|||||||:|||||||||||||:|||||||
Db         241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
              |||||||||||||:|
Db         301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         317 ------------------------------------------------------------ 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         317 ------------------------------------------------------------ 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
              |:|||:|||||||||||||:|||||||:|:|||||||||||||||||:|||||||||||
Db         458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:|||||||:|||||||||||||||||:|:||:|||||||||||||||||||||
Db         518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db         578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
```

```
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
Db       1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
Db       1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
Db       1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
Db       1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
Db       1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
Db       1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
Db       1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
Db       1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
Db       1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
Db       1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
Db       1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
Db       1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
Db       1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006
```

RESULT 102
AX027473
LOCUS AX027473 2006 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 23 from Patent FR2788784.
ACCESSION AX027473
VERSION AX027473.1 GI:10188437
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patent: FR 2788784-A 23 28-JUL-2000;
BIO MERIEUX (FR)
FEATURES Location/Qualifiers
source 1. .2006
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 24.7%; Score 1848.6; DB 1; Length 2006;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1949; Conservative 56; Mismatches 1; Indels 206; Gaps 3;

```
Qy       3957 ATGCAGTGGTCAGTGATAATGGAATATTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db          1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db         61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db        121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db        181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db        241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
Db        301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db        317 ---------------------------------------------------------- 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db        317 ---------------------------------------------------------- 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
Db        317 -----------------------------------AAGGGAAAATACTTTTGCCTG 337
```

```
Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
             |:|||:|||||||||||||||:|||||||:|:||||||||||||||||:||||||||||||
Db        458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:|||||||:|||||||||||||||||:|:||:|||||||||||||||||||||
Db        518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
             |||||||||||||||||||||||::||||||||||||||||||||||||||:||||||||||
Db       1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db       1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||||||||||||||:||:||||:|||:|||||:|||||||||
Db       1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             ||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||
Db       1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006
```

```
RESULT 103
AF072503
LOCUS       AF072503               2006 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
ACCESSION   AF072503
VERSION     AF072503.1  GI:4262287
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 2006)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.Pi22"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
```

```
     misc_feature    1. .2006
                     /note="corresponds to pol and env"
     repeat_region   1. .2006
                     /rpt_family="HERV-W"
     misc_feature    1420. .>2006
                     /note="corresponds to env gene"

  Query Match             24.7%;  Score 1848.6;  DB 1;  Length 2006;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1949;  Conservative   56;  Mismatches    1;  Indels  206;  Gaps    3;

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db             1 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 60

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
             ||:||||:|||||||||||||||:|||:|:|||||||||||||||||||||:||||||||
Db            61 CTCAGCTAGCAGAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAAAGGG 120

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
             :|||||||||||||||||||:|:|||||:|||||||| ||||||||||::|||||||
Db            121 CAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 180

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
             |:|:|||||||||||||||||||:|||:||||||||||||:||||||||||||||||||
Db            181 GGAAAGAAAGGGAATTCCTAACTTCTGAGAGAACACCTATCAAACATCAGGAAGCCATTA 240

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
             |||:|||||||:||||:|||||||||||||:||||||||:|||||||||||||:||||||
Db            241 GGAAATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTC 300

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
             |||||||||||||:|
Db            301 ATCANAAAGGAAAGGA---------------------------------------------- 316

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376

Db            317 ---------------------------------------------------------- 316

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436

Db            317 ---------------------------------------------------------- 316

Qy       4437 GAGGCAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
                                                 ||||||||||||||||||||||
Db            317 ----------------------------------------AAGGGAAAATACTTTTGCCTG 337

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            338 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 397

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db            398 GCACCCATCAAATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 457

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCAAGCTC 4676
             |:|||:|||||||||||||:|||||||:|:||||||||||||||:||||||||||||||
Db            458 AAATATTCAGGGCCTGTGAATTGTGCCAAAAAAATAATCCCCTGCCTCATCGCCAAGCTC 517

Qy       4677 CTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCA 4736
             ||||||||:|||||||:|||||||||||||||||:|:||:|||||||||||||||||||
Db            518 CTTCAGGAAAACAAAAAACAGGCCATTACCCTGAAAAAAACTGGCAACTGATTTTACCCA 577

Qy       4737 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCACGGGTTG 4796
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db            578 CAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATACTTTCACGGGTTG 637

Qy       4797 GGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 4856
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            638 GGCAAAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCA 697

Qy       4857 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db             698 TGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTT 757

Qy       4917 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             758 CCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACACTG 817

Qy       4977 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACA 5036
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db             818 CGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAATACTCAAAGGACA 877

Qy       5037 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTATAGCCTT 5096
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db             878 TCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTT 937

Qy       5097 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             938 AAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGG 997

Qy       5157 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             998 AAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAGTTGCAGAC 1057

Qy       5217 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1058 ATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGA 1117

Qy       5277 GAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCT 5336
             ||||||||||||||||||||||||||||::|||||||||||||||||||||||:||||
Db            1118 GAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCTCTCT 1177

Qy       5337 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db            1178 AATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATTTACCCCA 1237

Qy       5397 ACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGG 5456
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db            1238 ACTGCGGTTAAAGTGGCTGGAGTG--GTCTTGGATACATCACACTTGAGTCAAATCCTGG 1295

Qy       5457 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1296 ATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTA 1355

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1356 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 1415

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||||||||||||:||:||||:|||:||||||:|||||||||
Db            1416 CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCT 1475

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             ||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||
Db            1476 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATG 1535

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db            1536 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAC 1595

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1596 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 1655

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1656 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 1715

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1716 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 1775

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db            1776 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGT 1835
```

```
Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1836 ACATGGCACCTCTAG-CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 1894

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 1954

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCA 6168
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       1955 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAAGCCA 2006


RESULT 104
AF513360
LOCUS       AF513360               1860 bp    mRNA    linear   PRI 05-JUN-2002
DEFINITION  Homo sapiens enverin mRNA, complete cds.
ACCESSION   AF513360
VERSION     AF513360.1  GI:21326140
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1860)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7q and
            its env transcripts
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 312, 857-863 (1998)
REFERENCE   2  (bases 1 to 1860)
  AUTHORS   Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
            Rieger,F.
  TITLE     The HERV-W/7q family in the human genome. Potential for protein
            expression and gene regulation
  JOURNAL   Cell Mol. Biol. 48 (2), 213-217 (2002)
   PUBMED   11990458
REFERENCE   3  (bases 1 to 1860)
  AUTHORS   Alliel,P.M., Goudou,D., Perin,J.P. and Rieger,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-2002) U-488, INSERM, 80, rue du General Leclerc,
            Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .1860
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21"
                     /tissue_type="placenta"
     CDS             59. .1675
                     /codon_start=1
                     /product="enverin"
                     /protein_id="AAM47599.1"
                     /db_xref="GI:21326141"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"

  Query Match             24.5%;  Score 1834.3;  DB 1;  Length 1860;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches 1828;  Conservative   30;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5522 TTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCA 5581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 TTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCA 59

Qy       5582 TGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTG 5641
              |||:|||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||
Db         60 TGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTG 119

Qy       5642 CACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAA 5701
              ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db        120 CACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAA 179

Qy       5702 TGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCA 5761
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        180 TGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCA 239

Qy       5762 CCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 5821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 299

Qy       5822 ATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTG 5881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 ATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTG 359

Qy       5882 GAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTC 5941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTC 419

Qy       5942 AAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATG 6001
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        420 AAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATG 479

Qy       6002 GCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCC 6061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCC 539

Qy       6062 ATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCC 6121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCC 599

Qy       6122 AAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCC 6181
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        600 AAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCC 659

Qy       6182 CTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAG 6241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAG 719

Qy       6242 GACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTA 6301
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        720 GACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTA 779

Qy       6302 GCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAA 6361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAA 839

Qy       6362 TAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGA 6421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGA 899

Qy       6422 ATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACA 6481
              ||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||
Db        900 ATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACA 959

Qy       6482 CTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTC 6541
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        960 CTGAACAAGATTTATACAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTC 1019

Qy       6542 TTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCA 6601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 TTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCA 1079

Qy       6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 6661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1080 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 1139

Qy       6662 TCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTC 6721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 TCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTC 1199

Qy       6722 RAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAG 6781
             :|||||||||||||||||||:|||||||||||:|||||||||||||||||||||||||||
Db       1200 AAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAG 1259

Qy       6782 GGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAA 6841
             |||||||||||||||| ||||||||||||||||||||||||||||||||:||||||
Db       1260 GGGAAGAATGCTGTTACTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAA 1319

Qy       6842 TTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCC 6901
             |||:||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db       1320 TTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCC 1379

Qy       6902 TCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGC 6961
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGC 1439

Qy       6962 TACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCG 7021
             |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       1440 TACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCG 1499

Qy       7022 AAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 7081
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 AAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAC 1559

Qy       7082 CCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 7141
             |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 CCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTG 1619

Qy       7142 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 7201
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCG 1679

Qy       7202 GTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 7261
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAG 1739

Qy       7262 ACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGA 7321
             ||||||||||||||||||||||||||||:|||||||||:|||||||||:|||||||||||||
Db       1740 ACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGA 1799

Qy       7322 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGC 7381
             |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1800 CCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGC 1859

Qy       7382 T 7382
             |
Db       1860 T 1860
```

```
RESULT 105
AX478550
LOCUS       AX478550               2074 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 4 from Patent WO0248362.
ACCESSION   AX478550
VERSION     AX478550.1  GI:22217319
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ramkumar,J. and Arvizu,C.
  TITLE     Embryogenesis associated proteins
  JOURNAL   Patent: WO 0248362-A 4 20-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .2074
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7477736CB1"

  Query Match             24.2%;  Score 1812.4;  DB 1;  Length 2074;
  Best Local Similarity   91.7%;  Pred. No. 0;
  Matches 1904;  Conservative   33;  Mismatches  125;  Indels   14;  Gaps    3;

Qy       5517 GAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATC 5576
             |||||| || |||||||||||||| ||||||||  |||||||||||||||||||| |||||
Db          1 GAGGATCTGTGCCTGCTCTTCAAGCAACAACCGTGAGGAAAGTAACTAAAATCGTAAAT- 59

Qy       5577 CCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCT 5636
             ||||||||:|||||||||||  ||||||||||||:||:||||:| |:|||  :||||| |||
Db         60 CCCCATGGCCCTCCCTTATTGTATTTTTCTCTTTACTGTTCTCTCACCACCTTTCAGTCT 119

Qy       5637 CACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATG 5696
             |||||||||| ||||||||| ||||| ||||||||||||||||:|||:|| ||||||||||
Db         120 CACTGCACCCTCTCCATGCCACTGTAGGACCAGTAGCTCCCCTTACCAAGCGTTTCTATG 179

Qy       5697 GAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAC 5756
             |||||||| ||||||| || |||||||||||||||||||||||| | |:||||||||
Db         180 GAGAATGCGGCGTCCCAGACATATTGATGCCCCATCGTATAGGAGTTTATCTAAGGGAAA 239

Qy       5757 CCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTG 5816
             |||| ||||||| |||||||||||||||||| |||||||||| ||||||||||||||||||
Db         240 CCCCGCCTTCACCGCCCACACCCATATGCCCCACAACTGCTATAACTCTGCCACTCTTTG 299

Qy       5817 CATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGG 5876
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         300 TATGCATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGG 359

Qy       5877 ACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGG 5936
             |||||||| |||| |||||||||||||||||||| || ||||||||||||||||||||||
Db         360 ACTTGGAGCCACTATCTGTTGGACTTACTTCACCCATACCGGTATGTCTGATGGGGGTGG 419

Qy       5937 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGT 5996
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||:||||
Db         420 AGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGGAAGTAATCTCCCAACTCACCCGGGT 479

Qy       5997 ACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 6056
             |||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         480 ACATAGCACCCCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCG 539

Qy       6057 TACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 6116
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         540 TACCCATACTCACCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTC 599

Qy       6117 GGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTC 6176
             |||||||||||||||||||||||||||||| |||||||||||| |||||:|||||| ||||
Db         600 GGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTC 659

Qy       6177 AATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTT 6236
             | |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db         660 AGTCCCTGTACCTGAACAATGGAATAACTTCAGCACAGAAATAAACACCACTTCCATTTT 719

Qy       6237 AGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAA 6296
             |||||||||||||||||||||:|||||||||| ||||||||||||||||||||||||||||
Db         720 AGTAGGACCTCTTGTTTCCAATCTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAA 779

Qy       6297 ATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAC 6356
             |||||||||||||| |   |||||||||||||| |||||||| ||||||||||||||||||
Db         780 ATTTAGCAATACTATAGCAACAACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCAC 839

Qy       6357 ACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTG 6416
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         840 ACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTG 899

Qy       6417 TTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCAT 6476
             |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db         900 TTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCAT 959
```

```
Qy       6477 CTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACC 6536
Db        960 CTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACC 1019

Qy       6537 CATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGG 6596
Db       1020 CATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAG 1079

Qy       6597 TATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGA 6656
Db       1080 TATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGA 1139

Qy       6657 ACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGT 6716
Db       1140 ACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGT 1199

Qy       6717 CCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATT 6776
Db       1200 CCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATT 1259

Qy       6777 TTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRA 6836
Db       1260 TTTAGGAGAAGAATGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAA 1319

Qy       6837 AGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGG 6896
Db       1320 AGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACACTGGGG 1379

Qy       6897 CCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAAT 6956
Db       1380 CCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAAT 1439

Qy       6957 ATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAG 7016
Db       1440 ATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAG 1499

Qy       7017 AATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCCAAGAC 7064
Db       1500 AATTGAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGAGCCCCAGATGCAGTCCATGAC 1559

Qy       7065 TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACAT 7124
Db       1560 TAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACAT 1619

Qy       7125 CAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGC 7184
Db       1620 CGAAGGCACTCCTCCCAAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGC 1679

Qy       7185 AGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTG 7244
Db       1680 AGGAAGCAGTTAGAGCGGTCGTCAGTCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTG 1739

Qy       7245 AGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAG 7304
Db       1740 AGAGGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTAAGAATCCCTAAG 1799

Qy       7305 CCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACAC 7364
Db       1800 CCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAC 1859

Qy       7365 CTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATA 7424
Db       1860 CCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATA 1919

Qy       7425 GCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAA 7484
Db       1920 GCCAATCATCTATTGCCTGAGAGCACAGTGGGAGGGACAAGGATTGCAATATAAACCCAG 1979

Qy       7485 GTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGT 7544
Db       1980 GCATTCGAGCCAGC-ANAGCAACCGCCTTTGGGTCCCTTCCCCTTGTATGGGAGCTCTGT 2038

Qy       7545 TTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
Db       2039 TTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 2074
```

```
RESULT 106
CQ719381
LOCUS       CQ719381               2055 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 5315 from Patent WO02068579.
ACCESSION   CQ719381
VERSION     CQ719381.1  GI:42280238
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 5315 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .2055
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             23.6%;  Score 1768.9;  DB 1;  Length 2055;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1873;  Conservative   34;  Mismatches  120;  Indels   25;  Gaps    4;
```

```
Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
Db          1 ACAACTGTGAGGAAAGTAAC--AAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTT 57

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
Db         58 TTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACTGCACCCCCTCCATGCTGCTGTA 117

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
Db        118 CAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGAACGCGGCTTCCTGGAAATATTG 177

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
Db        178 ATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCCACCTTCACTGCCCACACCCATA 237

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
Db        238 TGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 297

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
Db        298 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGCCACTGTCTGTTGGACTT 357

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
Db        358 ACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATTCAAGGTCAGGCAAGAGAAAAAC 417

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAG 6022
Db        418 AAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCCTAGCCCCTACAAAG 477

Qy       6023 GACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTAT 6082
Db        478 GACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTGAGCCTAT 537

Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGA 6142
Db        538 TTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCTACTAACTGTTGGA 597
```

```
Qy       6143 TATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACA 6202
Db        598 TGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATCCCTGTTCCTGAACAATGGAACA 657

Qy       6203 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
Db        658 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGG 717

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
Db        718 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAACCA 777

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
Db        778 GCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCAGGAA 837

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
Db        838 TATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTATGT 897

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
Db        898 GCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATC 957

Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
Db        958 ATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAAAGCAG 1017

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
Db        1018 GAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTCTACT 1077

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
Db        1078 ACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACCT 1137

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
Db        1138 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACT 1197

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATG 6802
Db        1198 TGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTAGGAGAAGAATGCTGTTATTATG 1257

Qy       6803 TTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTA 6862
Db        1258 TTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAATGTA 1317

Qy       6863 KAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGA 6922
Db        1318 GAGCAGAGGAGCTTCAAAACACCGAACACTGGGGCCTCCTCAGCCAATGGATGCCCTGGG 1377

Qy       6923 TTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA 6982
Db        1378 TTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTA 1437

Qy       6983 TCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------- 7035
Db        1438 TCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTACAAATGG 1497

Qy       7036 -----CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACC 7090
Db        1498 TTCTTCAAATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACGGACCCCTGGACC 1557

Qy       7091 GGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCT 7150
Db        1558 AGCCTGCTAGCCCTTGTTCCGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCT 1617

Qy       7151 CAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGC 7210
Db        1618 CAACTGCACAACCCCTACCATGCCCCATTTCAGCAGGAAGCAGGTAGAGCGGTCGTCAGT 1677

Qy       7211 CAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTA 7270
Db        1678 CAACCTCCCCAGCAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGACAGGACTA 1737

Qy       7271 GCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCA 7330
Db        1738 GCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCA 1797

Qy       7331 CCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAA 7390
Db        1798 CCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAA 1857

Qy       7391 TGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCAC 7450
Db        1858 TGCTAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCAC 1917

Qy       7451 AGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCC 7510
Db        1918 AGCAAGAGGGACAATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCT 1977

Qy       7511 CTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAAT 7570
Db        1978 CTTTGGGTCCCCTCCCTTTGTATGGGAGCCCTGT----------TTTCACTCTATTAAAT 2027

Qy       7571 CTTGCARCTGCR 7582
Db        2028 CTTGCAACTGCA 2039
```

```
RESULT 107
AF156963
LOCUS       AF156963               1617 bp    DNA     linear   PRI 06-JUN-2000
DEFINITION  Homo sapiens human endogenous retrovirus W envC7-1 envelope protein
            (env) gene, complete cds.
ACCESSION   AF156963
VERSION     AF156963.1  GI:8272467
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1617)
  AUTHORS   Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
            and Paranhos-Baccala,G.
  TITLE     Chromosomal distribution and coding capacity of the human
            endogenous retrovirus HERV-W family
  JOURNAL   AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)
   PUBMED   10826480
REFERENCE   2  (bases 1 to 1617)
  AUTHORS   Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
            and Paranhos-Baccala,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-1999) UMR103CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon 69364, France
FEATURES             Location/Qualifiers
     source          1. .1617
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /clone="envC7-1"
                     /note="human/rodent monochromosomal NIGMS somatic hybrid
                     panel n#2 (Coriell Institute, Camden, NJ, USA); clone
                     GM/NA10791 retaining human chromosome 7"
     gene            1. .1617
                     /gene="env"
     CDS             1. .1617
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="AAF74215.1"
                     /db_xref="GI:8272468"
```

```
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLRNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPNSAGSS"
     repeat_region   1. .1617
                     /rpt_family="human endogenous retrovirus W"
                     /rpt_type=dispersed
     primer_bind     1. .18
     primer_bind     1597. .1617

  Query Match             21.5%;  Score 1607;  DB 1;  Length 1617;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1592;  Conservative   25;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||:|||||||||||:||||||||||||||||||||||||||||
Db       1141 CGAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db       1201 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||||||:||||||:|||||||||||||||||||||||||||||||
Db       1261 ATTCGAGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       1381 CTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 7080
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA 1500

Qy       7081 CCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 7140
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCT 1560

Qy       7141 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 1617


RESULT 108
AF506835
LOCUS       AF506835               1624 bp    mRNA    linear   PRI 20-MAY-2002
DEFINITION  Homo sapiens enverin mRNA, partial cds.
ACCESSION   AF506835
VERSION     AF506835.1  GI:20978307
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1624)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. II. HERV-7q
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
```

```
  PUBMED   9835022
REFERENCE  2  (bases 1 to 1624)
  AUTHORS  Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
           Rieger,F.
  TITLE    The HERV-W/7q family in the human genome. Potential for protein
           expression and gene regulation
  JOURNAL  Cell. Mol. Biol. 48, 213-217 (2002)
REFERENCE  3  (bases 1 to 1624)
  AUTHORS  Goudou,D., Perin,J.P., Rieger,F., Robert,B. and Alliel,P.M.
  TITLE    Direct Submission
  JOURNAL  Submitted (29-APR-2002) U-488, INSERM, 80, rue du General Leclerc,
           Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
     CDS             25. .>1624
                     /codon_start=1
                     /product="enverin"
                     /protein_id="AAM33413.1"
                     /db_xref="GI:20978308"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMTIYTEQDLYNYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILP
                     FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
                     RSDVNDIKGTPPEEISAAQPLLRPN"

  Query Match             21.4%;  Score 1600.3;  DB 1;  Length 1624;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 1597;  Conservative   25;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       5556 AAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGT 5615
             |||||||||||||||||| ||||||||:|||||||||||||||||||||:||:|||
Db          1 AAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGT 59

Qy       5616 TSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTC 5675
             |:|||:|||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db         60 TCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTC 119

Qy       5676 CCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTA 5735
             ||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTA 179

Qy       5736 TAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTG 5795
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTG 239

Qy       5796 CTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGAT 5855
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGAT 299

Qy       5856 TAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAAC 5915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAAC 359

Qy       5916 TGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGT 5975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 TGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGT 419

Qy       5976 AATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTC 6035
             ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        420 AATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTC 479

Qy       6036 AAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCT 6095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCT 539

Qy       6096 CACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCT 6155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCT 599

Qy       6156 GAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGA 6215
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGA 659

Qy       6216 AATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATAC 6275
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        660 AATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCATAC 719

Qy       6276 CTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCAT 6335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCAT 779

Qy       6336 CAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTG 6395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTG 839

Qy       6396 TGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATT 6455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATT 899

Qy       6456 CTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAA 6515
             ||||||||||||:|||:||||||||||||||||||||||||| ||||||||||||||||
Db        900 CTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCATATCTAA 959

Qy       6516 GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGC 6575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGC 1019

Qy       6576 ACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCA 6635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 ACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCA 1079

Qy       6636 AGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACT 6695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 AGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACT 1139

Qy       6696 TAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGA 6755
             ||||||||||||||||||||||||||:|||||||||||||||||||:||||||||||||||
Db       1140 TAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGA 1199

Qy       6756 RAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGG 6815
             :|||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1200 AAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTACTATGTTAATCAATCCGG 1259

Qy       6816 AATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCT 6875
             ||||||||||||||||||:|||||||||:||||||||||||||||||:||||||:||||||
Db       1260 AATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAGGAGCT 1319

Qy       6876 TCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTT 6935
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       1320 TCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTT 1379

Qy       6936 AGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCT 6995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       1380 AGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCT 1439

Qy       6996 TGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCA 7055
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       1440 TGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCA 1499

Qy       7056 GTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGT 7115
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db       1500 GTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGT 1559

Qy       7116 TAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCC 7175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 TAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCC 1619
```

```
Qy       7176 CAATT 7180
              |||||
Db       1620 CAATT 1624


RESULT 109
CQ053593
LOCUS       CQ053593                1894 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 4413 from Patent WO0157270.
ACCESSION   CQ053593
VERSION     CQ053593.1  GI:41028099
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 4413 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN HBL100, SIGNAL = 6"

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
              | |:|||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
              |||||||||||||||||||              ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA-------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
              |||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
              |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
              :|||||  |||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
              ||||||| |||||||||| ||||   ||||||||:|||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
              |||||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
```

```
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
              |||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
              ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
              |||||||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
              |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
              |||||||||||||||||:||||||||||||||||||:||||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
              |||||| |||||||||| |||| |||||||||| ||||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
              | |||||||||:|||||||||| || ||||||| ||||||:|||| |:|||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
              |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGA 7051
              |||||| |||||||||||||| ||:|||||             |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
              ||||||||| || ||| ||||||| | |||||||||||||||||:||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
              |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
              |||||||||||||||||||||||||||||| ||| || |||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
              |||||||||||||||| |||||||||||||||||| || || |||||||||  |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
              ||||||||:|||| ||||:|||||||||||||| | |||||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
              ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664
```

```
Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| |||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||          |||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 110
CQ068741
LOCUS       CQ068741               1894 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 4541 from Patent WO0157278.
ACCESSION   CQ068741
VERSION     CQ068741.1  GI:41038610
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 4541 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN HELA, SIGNAL = 7.2"

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             ||||||||||||||||||||             ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| ||||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
              :||||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||  ||||    ||||||||:|||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||| ||||||| ||  ||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||:|||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| |||||||| |||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||:|||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
             ||||||  |||||||||||||||| ||:||||||            |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             |||||||| || ||| ||||||| | ||||||||||||||||:||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||||| ||| || |||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
```

```
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351

Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402

Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462

Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522

Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582

Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 111
CQ095807
LOCUS       CQ095807                1894 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4666 from Patent WO0157272.
ACCESSION   CQ095807
VERSION     CQ095807.1  GI:41064833
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 4666 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN PLACENTA, SIGNAL = 20"

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803

Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863

Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923

Db        121 GTTGTCCTGGAGGACTTGGA-------------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983

Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043

Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103

Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163

Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223

Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283

Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343

Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403

Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463

Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523

Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583

Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643

Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703

Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763

Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823

Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883

Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943

Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003

Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAAATGGAGCCCAAGA 7051

Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111

Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364
```

```
Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             ||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||| ||| || |||||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| |||:|||||||||||||| | ||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         ||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||         ||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

RESULT 112
CQ134553
LOCUS CQ134553 1894 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 4575 from Patent WO0157276.
ACCESSION CQ134553
VERSION CQ134553.1 GI:41091909
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 4575 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1894
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002346.1
EXPRESSED IN BONE MARROW, SIGNAL = 6.3"

Query Match 20.7%; Score 1551.5; DB 1; Length 1894;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1679; Conservative 22; Mismatches 112; Indels 47; Gaps 5;

```
Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             ||||||||||||||||||||             ||||| ||||||| || ||  ||||||
```

```
Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||| ||||||||  |||| ||||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
              :|||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||| |||| ||||||||:|||||||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

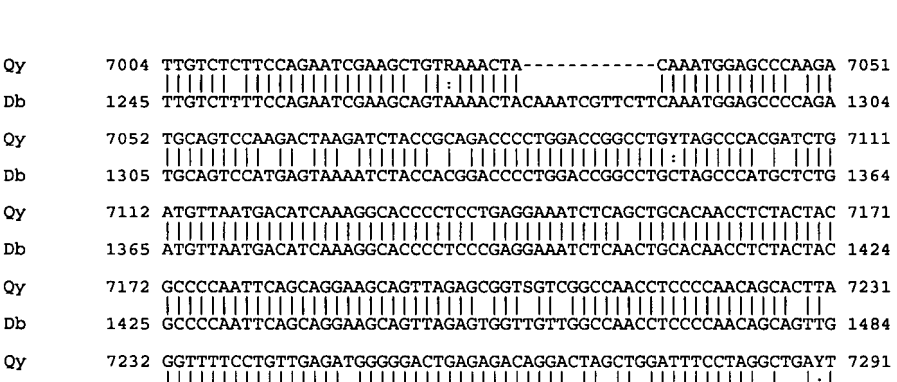
Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||:|||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| ||||||||| |||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:||||||||| || ||||||| |||||:|||| |:||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

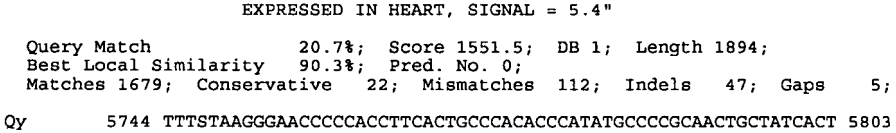
Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

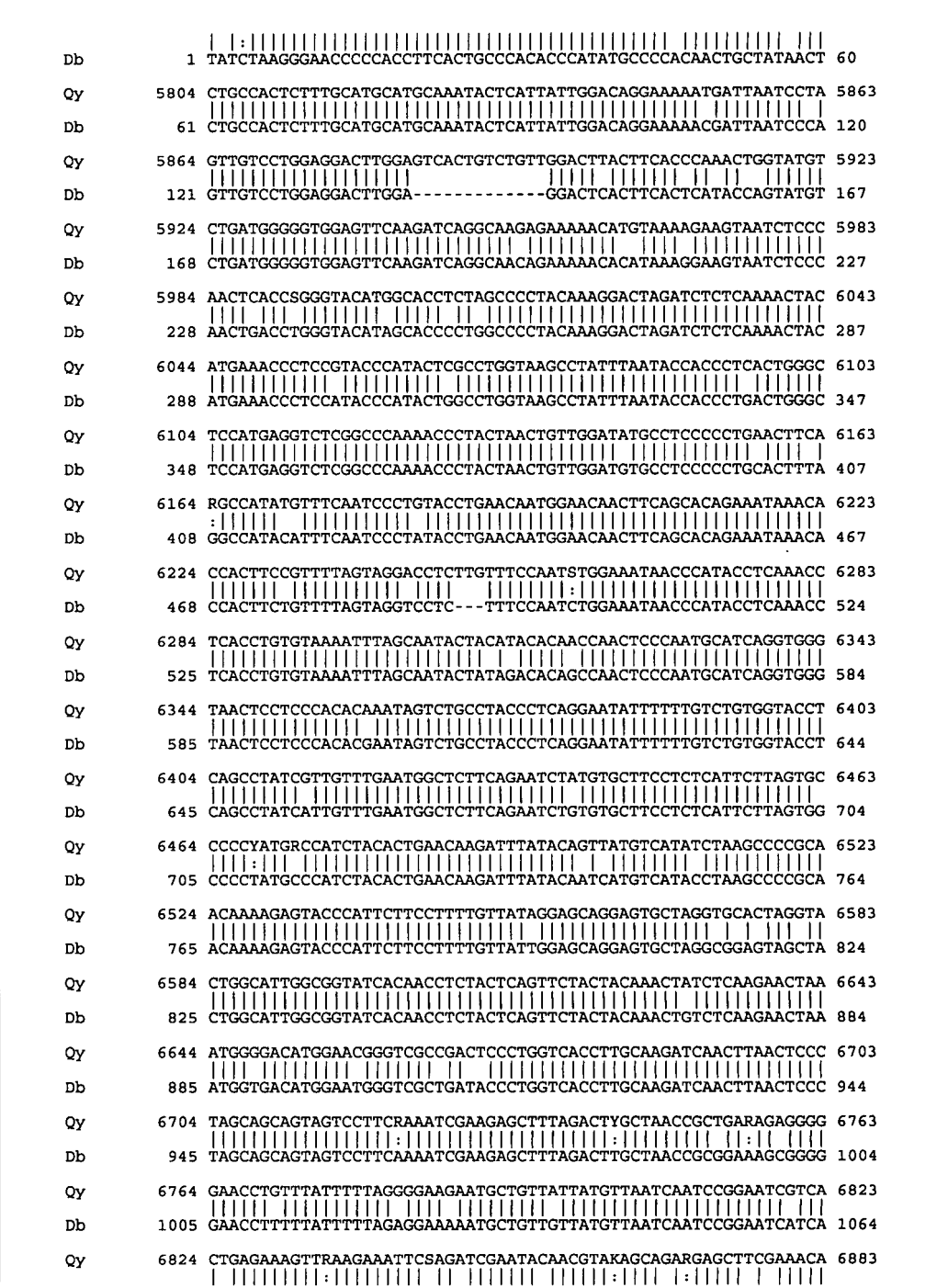
Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||:|||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244
```



RESULT 113
CQ173048
LOCUS CQ173048 1894 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 4444 from Patent WO0157274.
ACCESSION CQ173048
VERSION CQ173048.1 GI:41167784
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart
JOURNAL Patent: WO 0157274-A 4444 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1894
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002346.1
EXPRESSED IN HEART, SIGNAL = 5.4"

Query Match 20.7%; Score 1551.5; DB 1; Length 1894;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1679; Conservative 22; Mismatches 112; Indels 47; Gaps 5;





```
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943

Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003

Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051

Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111

Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171

Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231

Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291

Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351

Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402

Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462

Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522

Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582

Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834
```

```
RESULT 114
CQ202240
LOCUS       CQ202240               1894 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4517 from Patent WO0157271.
ACCESSION   CQ202240
VERSION     CQ202240.1  GI:41187918
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 4517 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN BT474, SIGNAL = 5.9"
```

Query Match 20.7%; Score 1551.5; DB 1; Length 1894;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1679; Conservative 22; Mismatches 112; Indels 47; Gaps 5;

```
Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803

Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863

Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923

Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983

Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043

Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103

Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163

Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223

Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283

Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343

Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403

Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463

Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523

Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583

Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643

Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703

Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
```

```
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823

Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883

Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943

Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003

Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051

Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111

Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171

Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231

Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291

Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351

Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402

Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462

Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522

Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582

Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 115
CQ217821
LOCUS       CQ217821                1894 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4660 from Patent WO0157273.
ACCESSION   CQ217821
VERSION     CQ217821.1  GI:41199145
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 4660 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 6.7"

  Query Match             20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803

Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863

Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923

Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983

Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043

Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103

Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163

Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223

Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283

Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343

Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

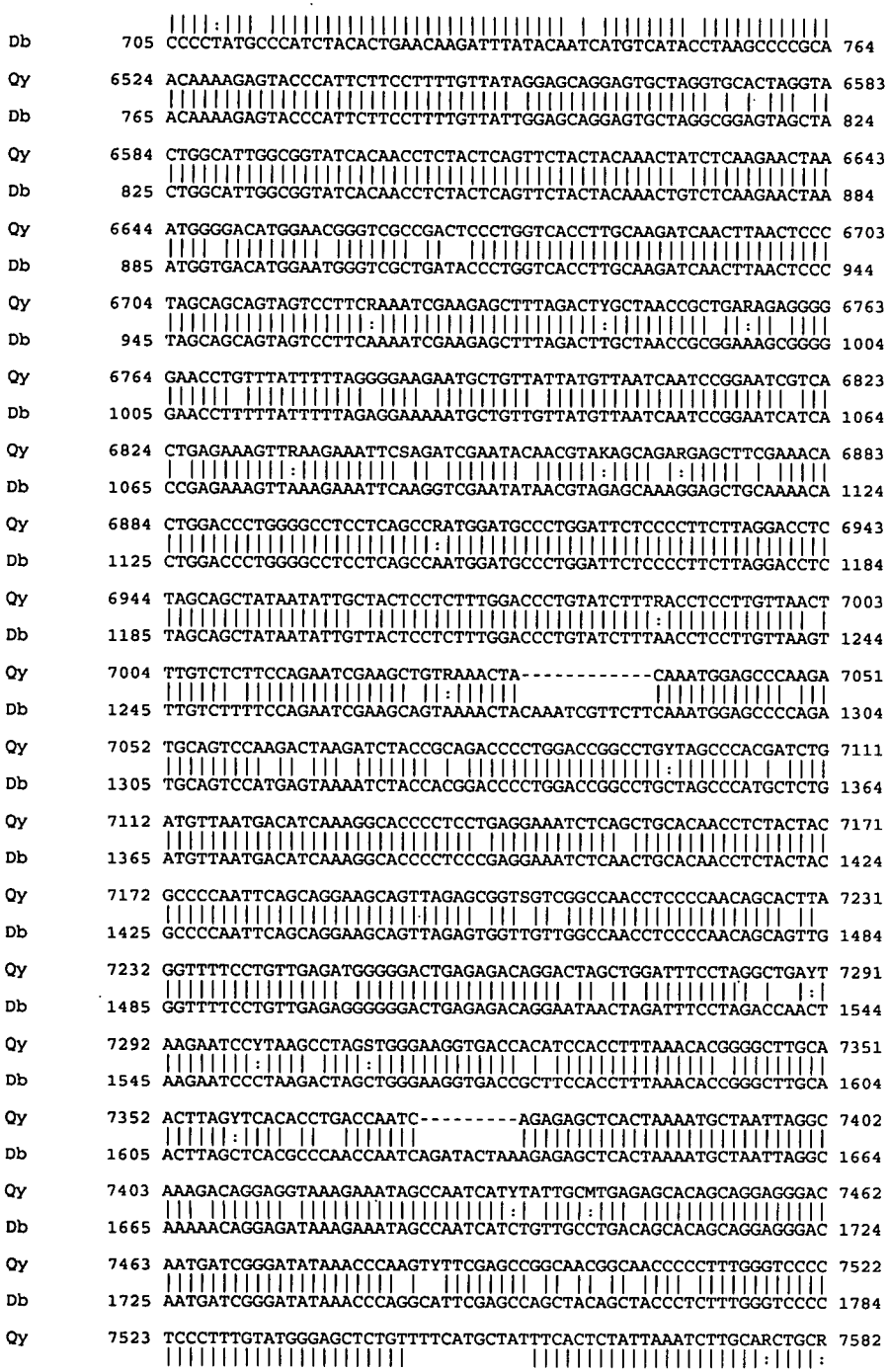
Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403

Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463

Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

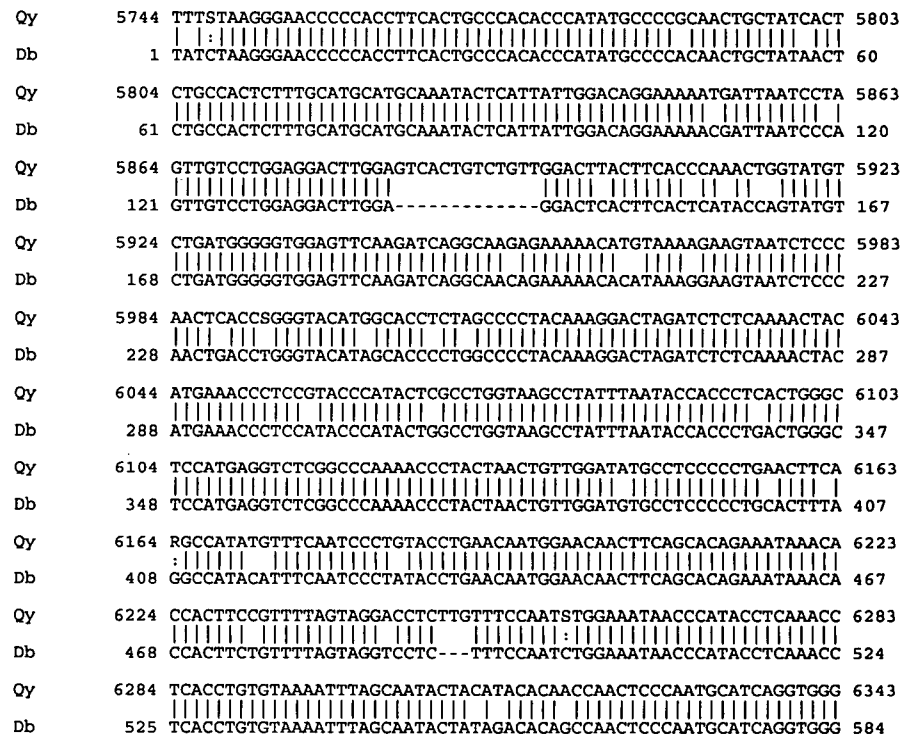
Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
```




Db 1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834

RESULT 116
CQ256381
LOCUS CQ256381 1894 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 4642 from Patent WO0157277.
ACCESSION CQ256381
VERSION CQ256381.1 GI:41228861
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 4642 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1894
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002346.1
EXPRESSED IN FETAL LIVER, SIGNAL = 4"

Query Match 20.7%; Score 1551.5; DB 1; Length 1894;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1679; Conservative 22; Mismatches 112; Indels 47; Gaps 5;



```
Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||| | |||||||| ||||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||| |||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             ||||||||||||||||:|||||||||||||||||:||||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||||| |||| ||||||||| |||||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||||:||||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||| ||:|||||| |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| ||||||| | ||||||||||||||||||:||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||| ||| || ||||||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| |||||||||||||||| | |||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| | ||||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          |||||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 117
CQ293475
LOCUS       CQ293475               1894 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4580 from Patent WO0186003.
ACCESSION   CQ293475
VERSION     CQ293475.1  GI:41254052
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 4580 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN LUNG, SIGNAL = 3.9"

  Query Match           20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches   112;  Indels   47;  Gaps    5;

Qy       5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db          1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy       5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db         61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy       5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||          ||||| ||||||| || ||  ||||||
Db        121 GTTGTCCTGGAGGACTTGGA----------GGACTCACTTCACTCATACCAGTATGT 167

Qy       5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||| |||||||||  |||| ||||||||||||
Db        168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy       5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||||
Db        228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy       6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||| |||||||
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
             :|||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467
```

```
Qy      6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||| ||||   ||||||||:|||||||||||||||||||||||||
Db       468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy      6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||
Db       525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy      6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy      6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db       645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy      6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||| | |||||||| ||||||||||||
Db       705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy      6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db       765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy      6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy      6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| ||||||||| ||||||| ||  ||||||||||||||||||||||||||||||||
Db       885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy      6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             ||||||||||||||||||:|||||||||||||||||||:|||||||| ||:|| ||||
Db       945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy      6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||| |||||||||| |||| ||||||||| |||||||||||||||||||||| |||
Db      1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy      6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:||||||||| || ||||||| ||||||:|||| |:||||| | |||||
Db      1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy      6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db      1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy      6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| ||||||||||||||||||||||:|||||||||||||| |
Db      1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy      7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
             |||||| |||||||||||||||| ||:|||||            |||||||||||| |||
Db      1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy      7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| |||||||| | |||||||||||||||||:|||||||| | ||||
Db      1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy      7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db      1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy      7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||| ||| || |||||||||||||||||||| ||
Db      1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy      7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||||| || || |||||||||| |  :|
Db      1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy      7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||||| | |||||||||||||| |||||||||
Db      1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy      7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         |||||||||||||||||||||||||||||
Db      1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

Qy      7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db      1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

Qy      7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db      1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy      7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||  |||||||||||||||||||||:||||:
Db      1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 118
CQ330413
LOCUS       CQ330413                1894 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4507 from Patent WO0157275.
ACCESSION   CQ330413
VERSION     CQ330413.1  GI:41278980
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 4507 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN BRAIN, SIGNAL = 6.2"

  Query Match           20.7%;  Score 1551.5;  DB 1;  Length 1894;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches 1679;  Conservative   22;  Mismatches  112;  Indels   47;  Gaps    5;

Qy      5744 TTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACT 5803
             | |:|||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db         1 TATCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCACAACTGCTATAACT 60

Qy      5804 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTA 5863
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        61 CTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAACGATTAATCCCA 120

Qy      5864 GTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGT 5923
             |||||||||||||||||||||              ||||| ||||||| || ||  ||||||
Db       121 GTTGTCCTGGAGGACTTGGA-------------GGACTCACTTCACTCATACCAGTATGT 167

Qy      5924 CTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCC 5983
             |||||||||||||||||||||||||||||||| ||||||||  |||| ||||||||||||
Db       168 CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAATCTCCC 227

Qy      5984 AACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 6043
             |||| ||| |||||||| ||||| || ||||||||||||||||||||||||||||||||
Db       228 AACTGACCTGGGTACATAGCACCCCTGGCCCCTACAAAGGACTAGATCTCTCAAAACTAC 287

Qy      6044 ATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGC 6103
             |||||||||||| |||||||||| |||||||||||||||||||||||||||| |||||||
```

```
Db        288 ATGAAACCCTCCATACCCATACTGGCCTGGTAAGCCTATTTAATACCACCCTGACTGGGC 347

Qy       6104 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCA 6163
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| |
Db        348 TCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTTA 407

Qy       6164 RGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 6223
              :||||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACTTCAGCACAGAAATAAACA 467

Qy       6224 CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACC 6283
             ||||||| |||||||||||  ||||    ||||||||:|||||||||||||||||||||||
Db        468 CCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATAACCCATACCTCAAACC 524

Qy       6284 TCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGG 6343
             |||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||||
Db        525 TCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCATCAGGTGGG 584

Qy       6344 TAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 6403
             |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCT 644

Qy       6404 CAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGC 6463
             |||||||||  |||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        645 CAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCATTCTTAGTGG 704

Qy       6464 CCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCA 6523
             ||||:||| |||||||||||||||||||||||||||||| | ||||||| ||||||||||||
Db        705 CCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAAGCCCCGCA 764

Qy       6524 ACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTA 6583
             |||||||||||||||||||||||||||||||||||||||||||||||||||| | | ||| ||
Db        765 ACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAGTAGCTA 824

Qy       6584 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAA 6643
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        825 CTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAACTAA 884

Qy       6644 ATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCC 6703
             |||| |||||||||  |||||||  ||  |||||||||||||||||||||||||||||||||
Db        885 ATGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTCCC 944

Qy       6704 TAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGG 6763
             |||||||||||||||||:|||||||||||||||||||:|||||||| ||:|| ||||
Db        945 TAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG 1004

Qy       6764 GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             ||||||  |||||||||||  ||||  |||||||||  |||||||||||||||||||||| |||
Db       1005 GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA 1064

Qy       6824 CTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACA 6883
             | |||||||||:|||||||||  || |||||||  ||||||:|||| |:|||||  | |||||
Db       1065 CCGAGAAAGTTAAAGAAATTCAAGGTCGAATATAACGTAGAGCAAAGGAGCTGCAAAACA 1124

Qy       6884 CTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 6943
             |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       1125 CTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTC 1184

Qy       6944 TAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACT 7003
             |||||||||||||||| |||||||||||||||||||||||||:||||||||||||||| |
Db       1185 TAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGT 1244

Qy       7004 TTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGA 7051
             |||||| ||||||||||||||| ||:||||||            |||||||||||| |||
Db       1245 TTGTCTTTTCCAGAATCGAAGCAGTAAAACTACAAATCGTTCTTCAAATGGAGCCCCAGA 1304

Qy       7052 TGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTG 7111
             ||||||||| || ||| ||||||| | |||||||||||||||||:||||||| | ||||
Db       1305 TGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTGCTAGCCCATGCTCTG 1364

Qy       7112 ATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTAC 7171
             |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db       1365 ATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCACAACCTCTACTAC 1424

Qy       7172 GCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTA 7231
             |||||||||||||||||||||||||||||| ||| || |||||||||||||||||||| ||
Db       1425 GCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCCAACAGCAGTTG 1484

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
             |||||||||||||||| |||||||||||||||||| || || |||||||||| |  |:|
Db       1485 GGTTTTCCTGTTGAGAGGGGGGACTGAGAGACAGGAATAACTAGATTTCCTAGACCAACT 1544

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
             ||||||||:|||| ||||:|||||||||||||| | |||||||||||||| |||||||||
Db       1545 AAGAATCCCTAAGACTAGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCA 1604

Qy       7352 ACTTAGYTCACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGC 7402
             ||||||:|||| ||  |||||||         ||||||||||||||||||||||||||||
Db       1605 ACTTAGCTCACGCCCAACCAATCAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGC 1664

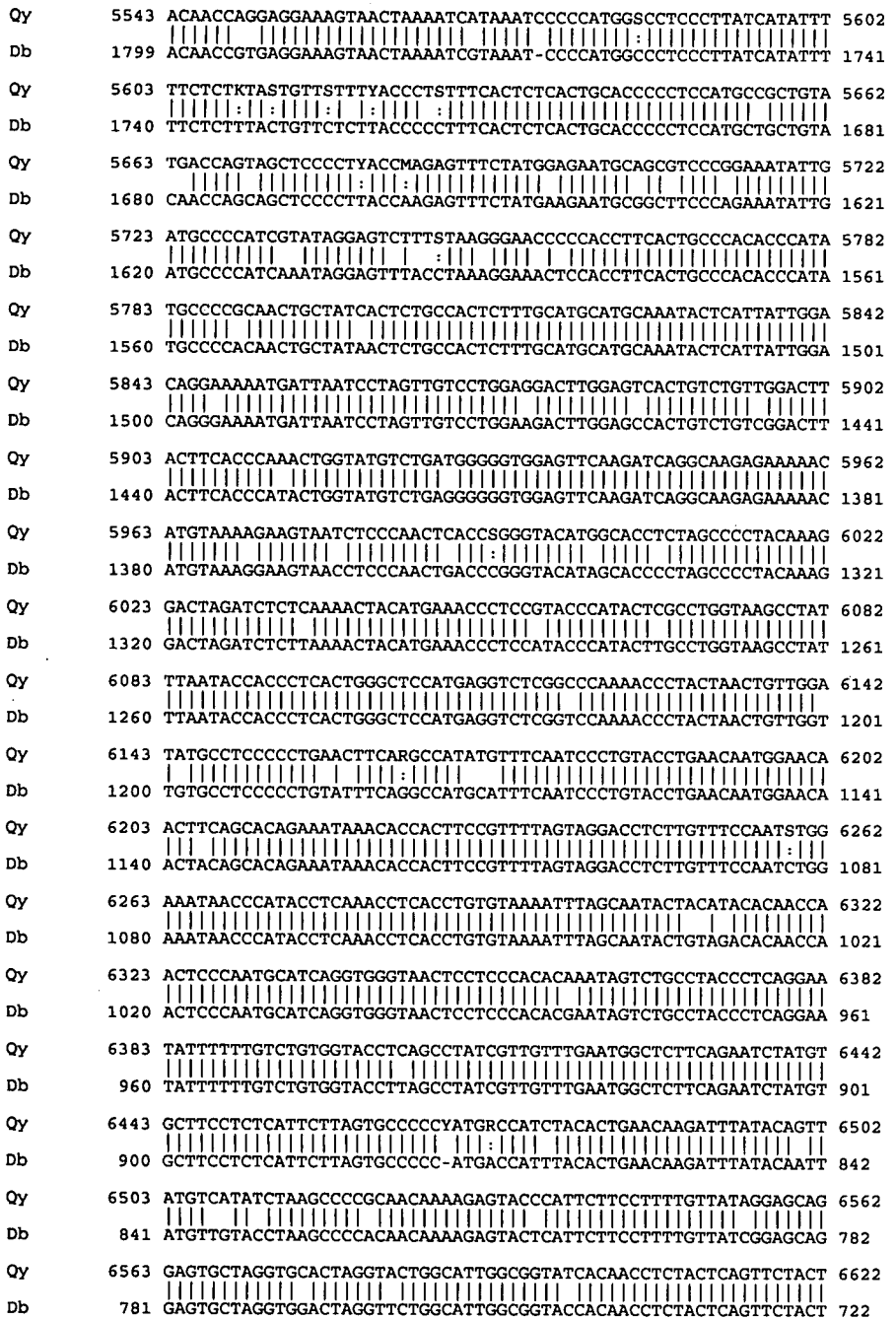
Qy       7403 AAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGAC 7462
             ||| ||||||| ||||||||||||||||||:| ||||:||| ||||||||||||||||||
Db       1665 AAAAACAGGAGATAAAGAAATAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGAC 1724

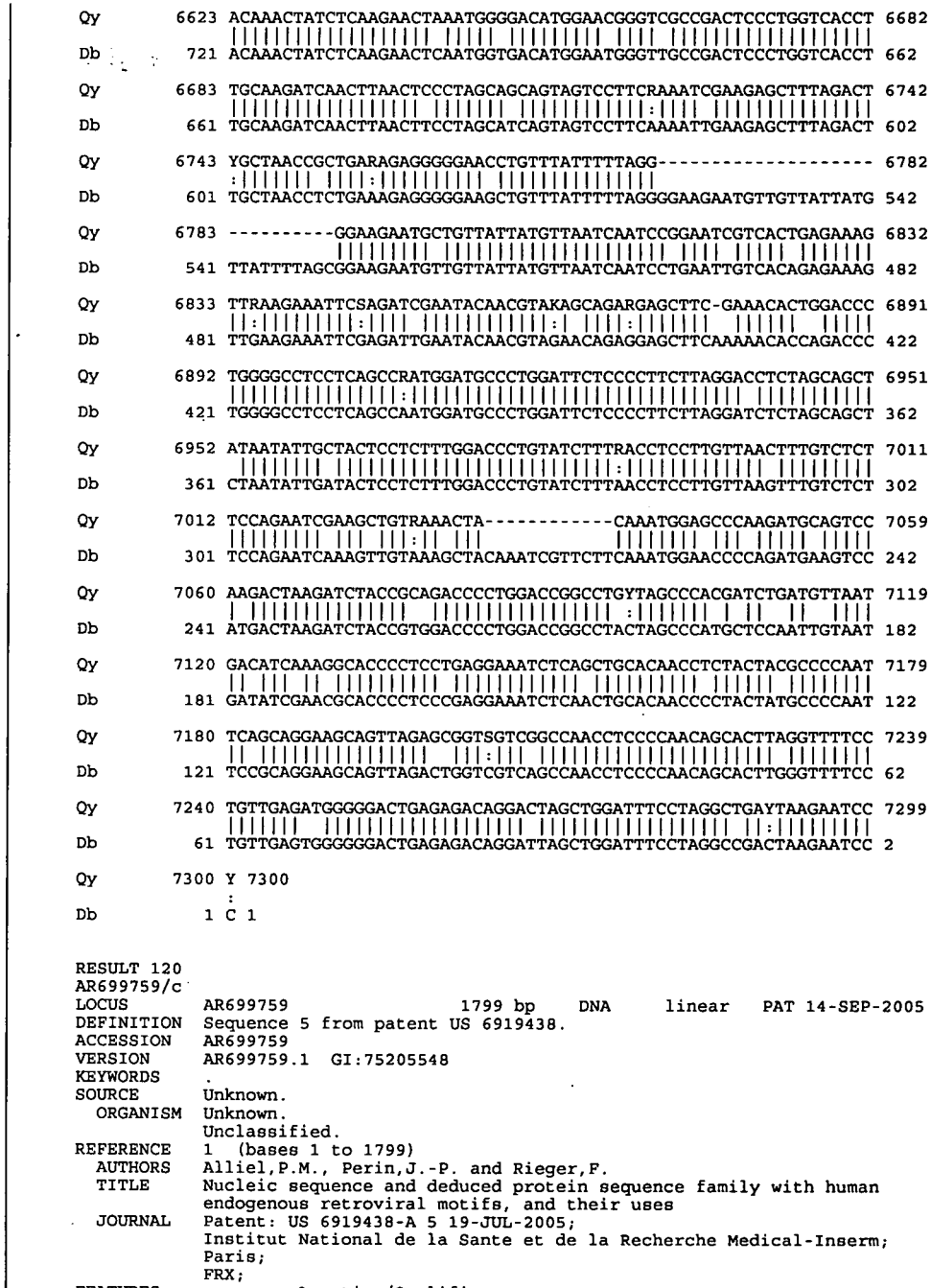
Qy       7463 AATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCC 7522
             |||||||||||||||||||| |  |||||||| || || || |||| ||||||||||||
Db       1725 AATGATCGGGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCC 1784

Qy       7523 TCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||||||||||          |||||||||||||||||||:||||:
Db       1785 TCCCTTTGTATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGCA 1834


RESULT 119
BD221810/c
LOCUS       BD221810                1799 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221810
VERSION     BD221810.1  GI:33031580
KEYWORDS    JP 2002518051-A/5.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1799)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 5 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002518051-A/5
            PD   25-JUN-2002
            PF   23-JUN-1999 JP 2000556036
            PR   23-JUN-1998 FR     98/07920
            PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
            C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
            A61P25/00,
            PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
            sequence and deduced protein sequence family with  CC
            human endogenous
            CC   retroviral motifs, and their uses
            FH   Key             Location/Qualifiers
            FT   source          1. .1799
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1799
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match           19.9%;  Score 1487.1;  DB 1;  Length 1799;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches 1613;  Conservative   27;  Mismatches  116;  Indels   45;  Gaps    5;
```

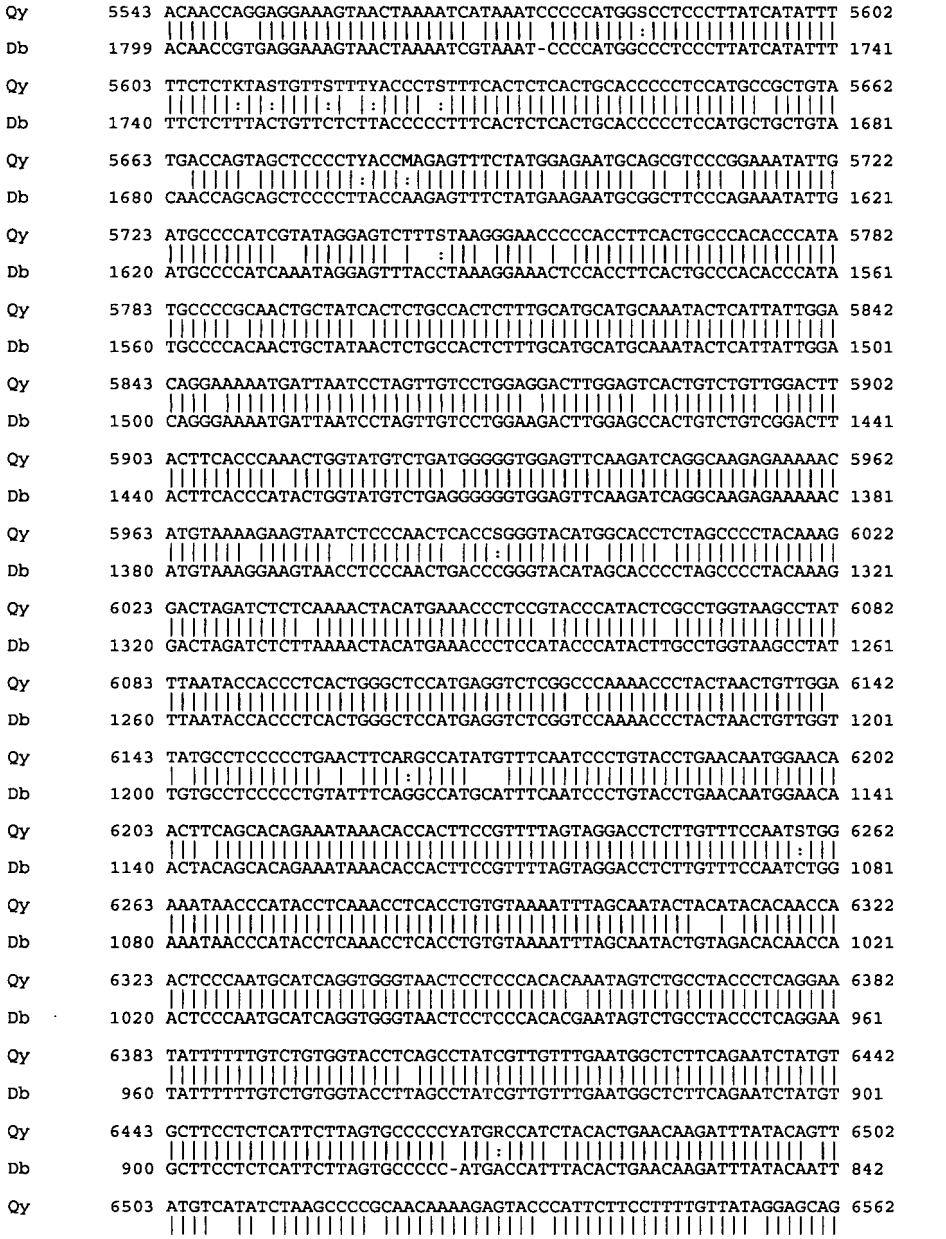


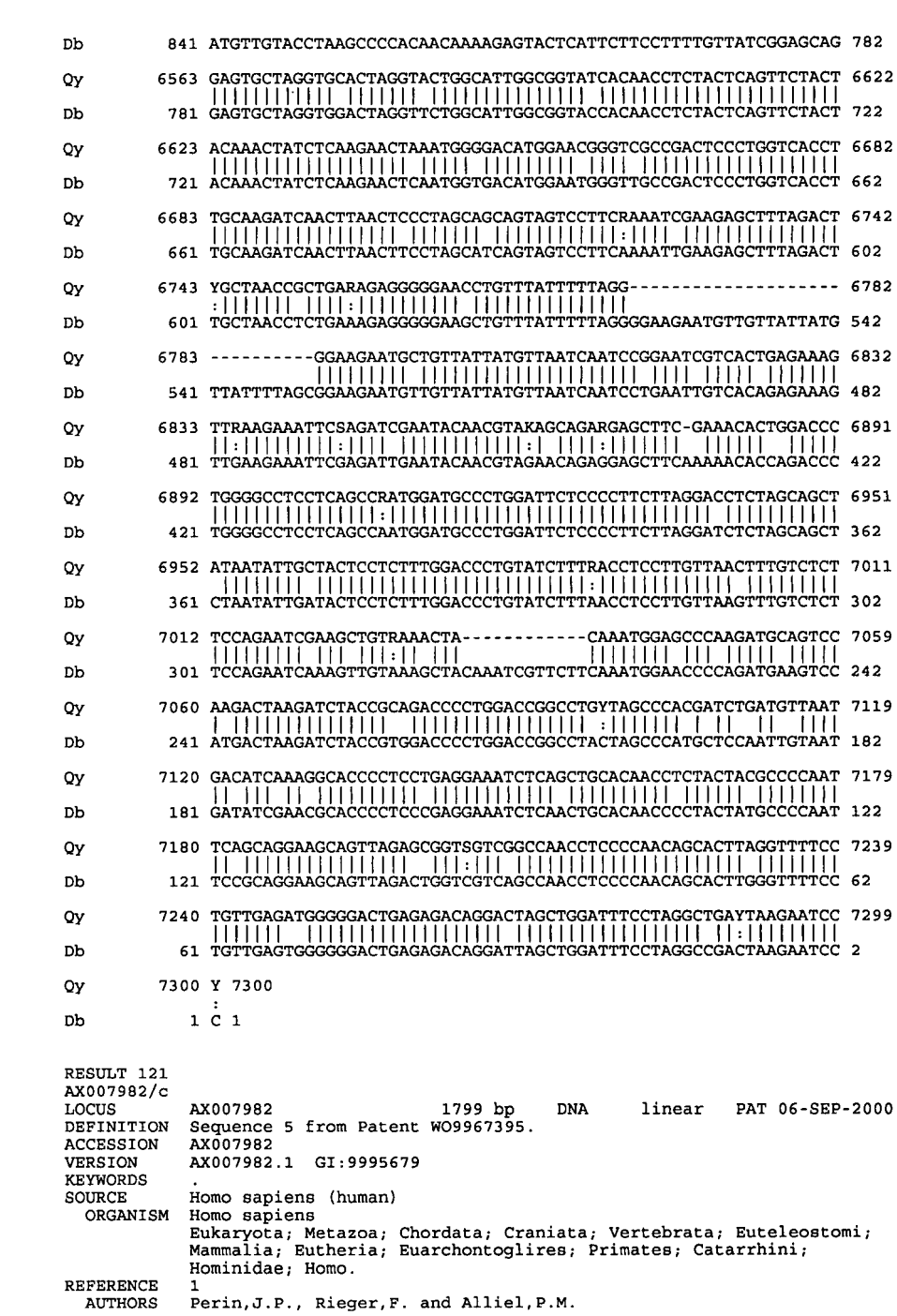


```
RESULT 120
AR699759/c
LOCUS       AR699759               1799 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 5 from patent US 6919438.
ACCESSION   AR699759
VERSION     AR699759.1  GI:75205548
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1799)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 5 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1799
```

/organism="unknown"
/mol_type="genomic DNA"

Query Match 19.9%; Score 1487.1; DB 1; Length 1799;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 1613; Conservative 27; Mismatches 116; Indels 45; Gaps 5;





RESULT 121
AX007982/c
LOCUS AX007982 1799 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent WO9967395.
ACCESSION AX007982
VERSION AX007982.1 GI:9995679
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Perin,J.P., Rieger,F. and Alliel,P.M.

```
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 5 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .1799
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             19.9%;  Score 1487.1;  DB 1;  Length 1799;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches 1613;  Conservative   27;  Mismatches  116;  Indels   45;  Gaps    5;

Qy       5543 ACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCATATTT 5602
             ||||||  |||||||||||||||||||| ||||| ||||||||:|||||||||||||||||
Db       1799 ACAACCGTGAGGAAAGTAACTAAAATCGTAAAT-CCCCATGGCCCTCCCTTATCATATTT 1741

Qy       5603 TTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGCTGTA 5662
             ||||||:||:||||:| |:|||| :|||||||||||||||||||||||||||| ||||||
Db       1740 TTCTCTTTACTGTTCTCTTACCCCCTTTCACTCTCACTGCACCCCCTCCATGCTGCTGTA 1681

Qy       5663 TGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATTG 5722
                ||||| |||||||||:|||:|||||||||||| ||||||| || |||| |||||||||
Db       1680 CAACCAGCAGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTTCCCAGAAATATTG 1621

Qy       5723 ATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACCCATA 5782
             ||||||||||  ||||||||  |  :||| |||| | ||||||||||||||||||||||||
Db       1620 ATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGCCCACACCCATA 1561

Qy       5783 TGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 5842
             ||||||  |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGA 1501

Qy       5843 CAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGTTGGACTT 5902
             ||||  |||||||||||||||||||||||||| |||||||| |||||||||  ||||||
Db       1500 CAGGGAAAATGATTAATCCTAGTTGTCCTGGAAGACTTGGAGCCACTGTCTGTCGGACTT 1441

Qy       5903 ACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 5962
             ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 ACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAC 1381

Qy       5963 ATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCCCTACAAAG 6022
             |||||||  |||||||  |||||||||  |||:||||||||  |||||  ||||||||||||||
Db       1380 ATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCTACAAAG 1321

Qy       6023 GACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTAT 6082
             ||||||||||||  |||||||||||||||||||||  ||||||||||  ||||||||||||||
Db       1320 GACTAGATCTCTTAAAACTACATGAAACCCTCCATACCCATACTTGCCTGGTAAGCCTAT 1261

Qy       6083 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGA 6142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 TTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAAACCCTACTAACTGTTGGT 1201

Qy       6143 TATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACA 6202
             |  ||||||||||||  | ||||:|||||    ||||||||||||||||||||||||||||||
Db       1200 TGTGCCTCCCCCTGTATTTCAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACA 1141

Qy       6203 ACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGG 6262
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1140 ACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGG 1081

Qy       6263 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCA 6322
             ||||||||||||||||||||||||||||||||||||||||||||||||  | |||||||||
Db       1080 AAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCA 1021

Qy       6323 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAA 6382
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1020 ACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAA 961

Qy       6383 TATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 6442
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
```

```
Db        960 TATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGT 901

Qy       6443 GCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTT 6502
             ||||||||||||||||||||||||||| |||:|||| |||||||||||||||||||||| ||
Db        900 GCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATT 842

Qy       6503 ATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAG 6562
             ||||  || |||||||||| |||||||||||||| |||||||||||||||||| |||||||
Db        841 ATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAG 782

Qy       6563 GAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACT 6622
             |||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||||||
Db        781 GAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACT 722

Qy       6623 ACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCT 6682
             |||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||||||
Db        721 ACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCT 662

Qy       6683 TGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACT 6742
             |||||||||||||||||||| ||||||| |||||||||||||:|||| ||||||||||||||
Db        661 TGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTAGACT 602

Qy       6743 YGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGG---------------------- 6782
             :|||||| ||||:|||||||||| ||||||||||||||
Db        601 TGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATG 542

Qy       6783 ----------GGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAG 6832
                       |||||||| |||||||||||||||||||| |||| ||||| |||||||
Db        541 TTATTTTAGCGGAAGAATGTTGTTATTATGTTAATCAATCCTGAATTGTCACAGAGAAAG 482

Qy       6833 TTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTC-GAAACACTGGACCC 6891
             ||:|||||||:|||| ||||||||||||:| ||||:|||||||  ||||| |||||
Db        481 TTGAAGAAATTCGAGATTGAATACAACGTAGAACAGAGGAGCTTCAAAAACACCAGACCC 422

Qy       6892 TGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCT 6951
             ||||||||||||||:|||||||||||||||||||||||||||||||| ||||||||||
Db        421 TGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGATCTCTAGCAGCT 362

Qy       6952 ATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCT 7011
             ||||||||  |||||||||||||||||||||||||||:|||||||||||||| ||||||||
Db        361 CTAATATTGATACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCT 302

Qy       7012 TCCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCC 7059
             |||||||||  ||| |||:|| |||              |||||||| ||| ||||| |||||
Db        301 TCCAGAATCAAAGTTGTAAAGCTACAAATCGTTCTTCAAATGGAACCCCAGATGAAGTCC 242

Qy       7060 AAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAAT 7119
             | ||||||||||||||  |||||||||||||||||| :|||||| | ||  ||  ||||
Db        241 ATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTACTAGCCCATGCTCCAATTGTAAT 182

Qy       7120 GACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAAT 7179
             || ||| || |||||||||| |||||||||||| |||||||||| |||||| ||||||||
Db        181 GATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTATGCCCCAAT 122

Qy       7180 TCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCC 7239
             || ||||||||||||||||  |||:||| |||||||||||||||||||| ||||||||
Db        121 TCCGCAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCCAACAGCACTTGGGTTTTCC 62

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||  |||||||||||||||||||| |||||||||||||||||| ||:|||||||||
Db          61 TGTTGAGTGGGGGGACTGAGAGACAGGATTAGCTGGATTTCCTAGGCCGACTAAGAATCC 2

Qy       7300 Y 7300
              :
Db          1 C 1
```

```
RESULT 122
BD196239
LOCUS       BD196239               1948 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196239
```

```
VERSION     BD196239.1  GI:33006009
KEYWORDS    JP 2002512530-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 5 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/5
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .1948
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||| |||||||:||||||
Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||
Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
             ||||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948


RESULT 123
BD267481
LOCUS       BD267481               1948 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267481
VERSION     BD267481.1  GI:33077249
```

```
KEYWORDS    JP 2002534980-A/22.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 22 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/22
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .1948
            FT                   /organism='Homo sapiens (human)'.
FEATURES                Location/Qualifiers
     source             1. .1948
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||
Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||
Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||||
Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
             ||||||||||||||||||||
Db       1930 CTGTTATTATGTTAATCAA 1948


RESULT 124
AX000960
LOCUS       AX000960               1948 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 5 from Patent WO9902696.
ACCESSION   AX000960
VERSION     AX000960.1  GI:7241202
KEYWORDS    .
SOURCE      unidentified
```

```
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 5 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db         671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:|||||:|||:||||||:|||||||||||||||||||||||||
Db         730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db         790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db         850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db        1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
             ||||||||||||||||||||
Db        1930 CTGTTATTATGTTAATCAA 1948


RESULT 125
AX027474
LOCUS       AX027474               1948 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 24 from Patent FR2788784.
ACCESSION   AX027474
VERSION     AX027474.1  GI:10188438
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 24 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
```

```
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471

Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531

Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591

Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651

Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711

Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771

Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831

Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891

Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951

Db       1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011

Db       1090 AAGAGAAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071

Db       1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131

Db       1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191

Db       1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251

Db       1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311

Db       1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371

Db       1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431

Db       1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491

Db       1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551

Db       1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611

Db       1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671

Db       1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731

Db       1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791

Db       1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810

Db       1930 CTGTTATTATGTTAATCAA 1948
```

```
RESULT 126
AF072505
LOCUS       AF072505                1948 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W envelope protein mRNA, partial
            cds.
ACCESSION   AF072505
VERSION     AF072505.1  GI:4262289
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 1948)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.24.4"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     repeat_region   1. .1948
                     /rpt_family="HERV-W"
     5'UTR           <1. .6
                     /note="R"
     5'UTR           7. .460
                     /note="U5"
     primer_bind     464. .481
                     /note="putative primer binding site for tRNA-W"
     misc_feature    491. .492
                     /note="putative splice junction"
     CDS             719. .>1948
                     /codon_start=1
```

```
                     /product="envelope protein"
                     /protein_id="AAD14545.1"
                     /db_xref="GI:4262290"
                     /translation="MALPYHIFLFTVVSPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLCKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSA
                     YRCLNGSSESMCFLSFLVPPMAIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALG
                     TGIGGITTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAE
                     RGGTCLFLGEECCYYVNQ"

  Query Match             19.2%;  Score 1440.3;  DB 1;  Length 1948;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1440;  Conservative   18;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5352 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTG 5411
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        491 GATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 550

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 610

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 670

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        671 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 729

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
             ||||||||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||
Db        730 TTATCATATTTTTCTCTTTACTGTTGTTTCACCCTCTTTCACTCTCACTGCACCCCCTCC 789

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
             ||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        790 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 849

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        850 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTGTAAGGGAACCCCCACCTTCACTGC 909

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 969

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 1029

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1030 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 1089

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
             ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        1090 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAG 1149

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1150 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 1209

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1210 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 1269

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        1270 TAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGA 1329

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1330 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 1389

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1390 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 1449

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1450 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 1509

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1510 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 1569

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGA 6491
             ||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db        1570 AGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGGCCATCTACACTGAACAAGA 1629

Qy       6492 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 6551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1630 TTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGT 1689

Qy       6552 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 6611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1690 TATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTAC 1749

Qy       6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 6671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1750 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTC 1809

Qy       6672 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        1810 CCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAG 1869

Qy       6732 AGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 6791
             ||||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db        1870 AGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATG 1929

Qy       6792 CTGTTATTATGTTAATCAA 6810
             ||||||||||||||||||||
Db        1930 CTGTTATTATGTTAATCAA 1948


RESULT 127
AF156961
LOCUS       AF156961               1923 bp    DNA     linear   PRI 06-JUN-2000
DEFINITION  Homo sapiens human endogenous retrovirus W gagC3.37 G gag (gag)
            gene, complete cds.
ACCESSION   AF156961
VERSION     AF156961.1  GI:8272463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1923)
  AUTHORS   Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
            and Paranhos-Baccala,G.
  TITLE     Chromosomal distribution and coding capacity of the human
            endogenous retrovirus HERV-W family
  JOURNAL   AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)
   PUBMED   10826480
REFERENCE   2  (bases 1 to 1923)
  AUTHORS   Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
            and Paranhos-Baccala,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-1999) UMR103CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon 69364, France
FEATURES             Location/Qualifiers
     source          1. .1923
                     /organism="Homo sapiens"
```

```
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="gagC3.37"
                     /note="human/rodent monochromosomal NIGMS somatic hybrid
                     panel n#2 (Coriell Institute, Camden, NJ, USA); from clone
                     GM/NA10253 retaining human chromosome 3"
     primer_bind     1. .20
     repeat_region   21. .1900
                     /rpt_family="human endogenous retrovirus W"
                     /rpt_type=dispersed
     gene            348. .1439
                     /gene="gag"
     CDS             348. .1439
                     /gene="gag"
                     /codon_start=1
                     /product="gag"
                     /protein_id="AAF74213.1"
                     /db_xref="GI:8272464"
                     /translation="MGNIPPKAKTPLRCILENWDQCDTQTLRKKRFIFFCSTAWPQYP
                     LQGRETWLPEGSINYNIILQLDLFCRKEGKWSEVPYVQTFFSLRDNSQLCKKCGLCPT
                     GSPQSPPPYPSVPPPTPSSTNKDPPLTQTVQKEIDKGVNNEPKSANIPRLCPLQAVRG
                     GEFGPARVPVPFSLSDLKQIKIDLGKFSDNPDGYIDVLQGLGQSFDLTWRDIMLLLNQ
                     TLTPNERSAAVTAAREFGDLWYLSQVNNRMTTEERTTPTGQQAVPSVDPHWDTESEHG
                     DWCHKHLLTCVLEGLRKTRKKPMNYSMMSTITQGKEENPTAFLDRLREALRKHTSLSP
                     DSIEGQLILKDKFITQSAADIRKNFKSLP"
     primer_bind     1901. .1923

  Query Match             19.2%;  Score 1434.3;  DB 1;  Length 1923;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches 1698;  Conservative   27;  Mismatches  148;  Indels   87;  Gaps   15;

Qy        237 TGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGC 296
              ||||||||||||||||||||||||||:||||||||||| ||||||||||||||||||||||
Db          1 TGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCTGCTCCCAATTGGGCTAAAGGC 60

Qy        297 TTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANT 356
              |||||||||| ||  |||||||||||||| |||||  | ||||||||||||||||||| |
Db         61 TTGCCATTGTTCCCACACGGCTAAGTGCCCGGGTTCATCCTAATTGAGCTGAACACTAGT 120

Qy        357 CACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTAC 416
              ||||||||||||||||||||||  ||||||||||||||||:| || ||| ||| ||
Db        121 CACTGGGTTCCATGGTTCTCTTCCATGACCCACGGCTTCTAATAGAGCTCTAATACTCAC 180

Qy        417 CACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGA 476
              ||||||||||||||||||||||||||||||||||||:||| |||| ||||||||||||
Db        181 CACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACCCCAGGTCAGAGA 239

Qy        477 ATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCAC 536
              | ||||:|||||||||| |||||||| ||||                    | || |
Db        240 ACACGAGGCTTGCCACCGTCTTGGAAGTGGCC--------------------CGCCGC 277

Qy        537 CATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGAC 596
              |||||||||||||||| |||||||||||||| ||:||||||||:||||: :| ||||
Db        278 CATCTTGGGAGCTCTGGGAGCAAGGACCCCCCAGTAACATTTTGGCAACCACAAAGGGAC 337

Qy        597 ATCCMAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGA 656
              |||:||| |||||||||| |||||| |||| |||||||||||||||| ||||||||||
Db        338 CTCCAAAGCGATGGGAAACATTCCCCCCAAGGCAAAAACGCCCCTAAGATGTATTCTGGA 397

Qy        657 RAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAG 716
              :||||||||:|||| |||| ||||||| |||||||||||||||| ||||||||||||||||
Db        398 GAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACGATTTATATTCTTCTGCAG 457

Qy        717 TGCCGCCTGGC----------------------------------ACTCCTGAGGGAAGTAT 744
              | |||||||||                                  ||||||||||||||||
Db        458 TACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTGGCTTCCTGAGGGAAGTAT 517

Qy        745 AAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAA---GGCAAATGGAGTGA 801
              |||||||||||| ||||||||||||||:|||| ||||||||    ||||||||||||||||
Db        518 AAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAAGGAGGGCAAATGGAGTGA 577

Qy        802 AGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTG 861
              |||||||||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db        578 AGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTG 637

Qy        862 TGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCAT-----CCCC 916
              || |||||||||||||||||||||| |||||||| |||||||| |||||| |     ||||
Db        638 TGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTACCCCAGCGTCCCCCCCCC 697

Qy        917 GACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGA 976
              |||||||||| ||::|||||||||||||||| |||||||| ||||||||||||||||||
Db        698 GACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAACGGTCCAAAAGGAGATAGA 757

Qy        977 CAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAG 1036
              |||| |||||||||| ||||||||||||||||||||||||| |||||| ||||| |||||
Db        758 CAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCGATTATGCCCCCT-CCAAG 816

Qy       1037 CAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACT 1096
              ||||| ||||| |||||||||||||||||||||||||| ||| |:||||::|||| ||||||
Db        817 CAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTACCTTTTTCTCTCTCAGACT 876

Qy       1097 TAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTG 1156
              |||||||||| |||| |||| ||||||||||||||||||:||||| ||||||||||:||
Db        877 TAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACCCTGACGGCTATATTGATG 936

Qy       1157 TTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTA 1216
              |||||||||||||||||||| |||||||||||||||||||||||| |||| ||| |||
Db        937 TTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAGATATA-ATGTTACTACTA 995

Qy       1217 AATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGC 1276
              |||||||||||||||||||||||||||||||||||| |  |||||||||| |||:|||||||
Db        996 AATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACTGCAGCCCGAGAGTTTGGC 1055

Qy       1277 GATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGA 1336
              ||||| |||||||||||||||||||   |   |||||||||||| | ||  | ||
Db       1056 GATCTTTGGTATCTCAGTCAGGTCAA--CAATAGGATGACAACAGAGGAAAGAACAA--- 1110

Qy       1337 TTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAG 1396
              ||||||||||||||:|||||||||||| ||:|||||||||||   ||| | |
Db       1111 -CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCATTGGG---ACACAGAATCA 1166

Qy       1397 TAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAA 1456
              ||||| |||||||||||  || |||||||||||||| ||||||:||||||| |||||
Db       1167 GAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGTGCTAGAAGGACTGAGGAA 1225

Qy       1457 AACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAA 1516
              |||||:|||| ||: |||:|||||||||||||||||||| |||||||| ||||| ||||
Db       1226 AACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTATAACACA-GGGAAAGGAA 1283

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
              |||||||||||||| |||||||| |||||||||||||||||||||||| | |||| |||
Db       1284 GAAAATCCTACTGCTTTTCTGGACAGACTAAGGGAGGCATTGAGGAAGCATACCTCCCTG 1343

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
              |||||||||||  |||||||||||||||||||  ||||||||||||||||||||||||||
Db       1344 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTCAGTCAGCT 1403

Qy       1637 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCC 1696
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db       1404 GCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGGAGCAGAACTTAGAAACCC 1463

Qy       1697 TATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGG 1756
              |||| ||||||||| |:|| |||||||||||||||||||||||||||||||| ||| ||
Db       1464 TATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGAAACGGG 1523

Qy       1757 ACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGTCATGACCCTCAGGCAAGT 1809
              |||||||||| ||||||||     | ||||  |||||||||| ||||||||||
Db       1524 ACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGTCATGGCCCTCAGGCAAGC 1583

Qy       1810 GGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTG 1869
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db       1584 AGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATCAAATGCCTAATAGGGCTGG 1643

Qy       1870 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCC 1929
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||||
```

```
Db       1644 CTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCCAAGTAGAAATAAGCCGCC 1703

Qy       1930 CCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACA 1989
             || | ||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db       1704 CCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATG 1763

Qy       1990 AAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCT 2049
             ||| |  | |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1764 AAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCC 1823

Qy       2050 GGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGA 2109
             |||||  ||||||| |||||||||||||||||||||||||||  ||||||||||||||||||
Db       1824 GGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGA 1883

Qy       2110 GGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGGTC 2146
              | |||||||  | |   ||||||||||||||||| || | |
Db       1884 GAGCCAGGAGGTTAACTGTCTCCTGGACACTGGCGCAGCC 1923


RESULT 128
BD252040
LOCUS       BD252040                1629 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Method for detecting superantigen activity in a biological sample.
ACCESSION   BD252040
VERSION     BD252040.1  GI:33061810
KEYWORDS    JP 2002539804-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1629)
  AUTHORS   Perron,H. and Lafont,M.
  TITLE     Method for detecting superantigen activity in a biological sample
  JOURNAL   Patent: JP 2002539804-A 1 26-NOV-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002539804-A/1
            PD   26-NOV-2002
            PF   20-MAR-2000 JP 2000607009
            PR   19-MAR-1999 FR     99/03622,28-OCT-1999 FR     99/13755 PI
            HERVE PERRON,MONIQUE LAFONT
            PC   C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K45/00,A61K48/00,
            PC   A61P25/00,
            PC   C07K14/47,C07K16/18,C12N7/00,C12Q1/02,C12Q1/68,G01N27/447, PC
             G01N33/15,
            PC
            G01N33/50,G01N33/569,G01N33/577//C12P21/08,(C12N7/00,C12R1:93), PC
             (C12Q1/02,C12R1:93),(C12Q1/68,C12R1:93),C12N15/00,G01N27/26, PC
            A61K37/02
            CC   Method for detecting superantigen activity in a biological  CC
                 sample
            FH   Key             Location/Qualifiers
            FT   source          1. .1629
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1629
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             19.1%;  Score 1431;  DB 1;  Length 1629;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches 1492;  Conservative   25;  Mismatches  100;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||| ||||||  |||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              | || || ||| |||||||||||||||||||| |||||||| | |:|||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||| |||||||||||||||||||||||||| ||  |||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||| |||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              |||||  |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||| |||||||||||||||||||||| ||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||||| |||||||||||| ||||||:|||||| |||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||| | |||||||||| |||||||||||||||||||||||| |||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              |||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
              |||||||||||||||||| | ||||| || |||||||| ||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
              |||||||||||||| |||||||||||||||  |||||||||||||||||||| |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
              |||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||| |
Db        1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
              ||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
              |:||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||||
Db        1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
```

```
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960

Db       1321 CTCAGCCAATGGATGCCCTGGACTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTT 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020

Db       1381 TTACTCCTCTTTGGACCCTGTATCTTCAACTTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068

Db       1441 GAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAACCCCAGATGCAGTCCATGACTAAA 1500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128

Db       1501 ATCTACCGTGGACCCCTGGACCGGCCTGCTAGACTATGCTCTGATGTTAATGACATTGAA 1560

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188

Db       1561 GTCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTACACTCCAATTCAGTAGGA 1620

Qy       7189 AGCAGTTAG 7197

Db       1621 AGCAGTTAG 1629


RESULT 129
AX036992
LOCUS       AX036992               1629 bp    DNA     linear   PAT 16-NOV-2000
DEFINITION  Sequence 1 from Patent FR2791140.
ACCESSION   AX036992
VERSION     AX036992.1  GI:11226420
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lafont,M. and Perron,H.
  JOURNAL   Patent: FR 2791140-A 1 22-SEP-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1629
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             19.1%;  Score 1431;  DB 1;  Length 1629;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches 1492;  Conservative   25;  Mismatches  100;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640

Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700

Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760

Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820

Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880

Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940

Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000

Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060

Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120

Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180

Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240

Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300

Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360

Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420

Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480

Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540

Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600

Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660

Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720

Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780

Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840

Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320
```

```
Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGACTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTT 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              |||||||||||||||||||||||| :|| |||||||||| ||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTCAACTTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTA----------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
             ||||||||:|| |||     ||||||| ||| ||||||||||| ||||||
Db       1441 GAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAACCCCAGATGCAGTCCATGACTAAA 1500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              ||||||||  ||||||||||||||||||:||| | | | ||||||||||||||||  ||
Db       1501 ATCTACCGTGGACCCCTGGACCGGCCTGCTAGACTATGCTCTGATGTTAATGACATTGAA 1560

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
              | ||||||||| |||||||||||| |||||||||| ||||||| | ||||||||| ||||
Db       1561 GTCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTACACTCCAATTCAGTAGGA 1620

Qy       7189 AGCAGTTAG 7197
              |||||||||
Db       1621 AGCAGTTAG 1629


RESULT 130
AF331500
LOCUS       AF331500               1629 bp    mRNA    linear   VRL 21-NOV-2001
DEFINITION  Multiple sclerosis associated retrovirus element clone pV14
            recombinant envelope protein (env) mRNA, complete cds.
ACCESSION   AF331500
VERSION     AF331500.1  GI:13310190
KEYWORDS    .
SOURCE      Multiple sclerosis associated retrovirus element
  ORGANISM  Multiple sclerosis associated retrovirus element
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 1629)
  AUTHORS   Perron,H., Jouvin-Marche,E., Michel,M., Ounanian-Paraz,A.,
            Camelo,S., Dumon,A., Jolivet-Reynaud,C., Marcel,F., Souillet,Y.,
            Borel,E., Gebuhrer,L., Santoro,L., Marcel,S., Seigneurin,J.M.,
            Marche,P.N. and Lafon,M.
  TITLE     Multiple sclerosis retrovirus particles and recombinant envelope
            trigger an abnormal immune response in vitro, by inducing
            polyclonal Vbeta16 T-lymphocyte activation
  JOURNAL   Virology 287 (2), 321-332 (2001)
   PUBMED   11531410
REFERENCE   2  (bases 1 to 1629)
  AUTHORS   Perron,H., Ounanian-Paraz,A., Marcel,F. and Souillet,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-2000) R&D, bioMerieux, Chemin de L'Orme, Marcy
            L'Etoile 69280, France
FEATURES             Location/Qualifiers
     source          1. .1629
                     /organism="Multiple sclerosis associated retrovirus
                     element"
                     /mol_type="mRNA"
                     /db_xref="taxon:89382"
                     /clone="pV14"
     gene            1. .1629
                     /gene="env"
     CDS             1. .1629
                     /gene="env"
                     /function="displaying superantigen-like properties causing
                     polyclonal activation of human Vbeta 16 T-cells (CD3+)
                     independent from the T-cell receptor antigen specificity"
                     /note="surface and transmembrane regions"
                     /codon_start=1
                     /product="recombinant envelope protein"
                     /protein_id="AAK18189.1"
                     /db_xref="GI:13310191"
                     /translation="MALPYHTFLFTVLLPPFALTAPPPCCCTTSSSPYQEFLWRTRLP
                     GNIDAPSYRSLSKGNSTFTAHTHMPRNCYNSATLCMHANTHYWTGKMINPSCPGGLGA
                     TVCWTYFTHTSMSDGGGIQGQAREKQVKEAISQLTRGHSTPSPYKGLVLSKLHETLRT
                     HTRLVSLFNTTLTRLHEVSAQNPTNCWMCLPLHFRPYISIPVPEQWNNFSTEINTTSV
                     LVGPLVSNLEITHTSNLTCVKFSNTIDTTSSQCIRWVTPPTRIVCLPSGIFFVCGTSA
                     YHCLNGSSESMCFLSFLVPPMTIYTEQDLYNHVVPKPHNKRVPILPFVIRAGVLGRLG
                     TGIGSITTSTQFYYKLSQEINGDMEQVTDSLVTLQDQLNSLAAVVLQNRRALDLLTAK
                     RGGTCLFLGEERCYYVNQSRIVTEKVKEIRDRIQCRAEELQNTERWGLLSQWMPWTLP
                     FLGPLAAIIFLLLFGPCIFNFLVKFVSSRIEAVKLQIVLQMEPQMQSMTKIYRGPLDR
                     PARLCSDVNDIEVTPPEEISTAQPLLHSNSVGSS"

  Query Match             19.1%;  Score 1431;  DB 1;  Length 1629;
  Best Local Similarity   91.6%;  Pred. No. 0;
  Matches 1492;  Conservative   25;  Mismatches  100;  Indels   12;  Gaps    1;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              |||||||||||||||| ||||||  |||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              | || || ||| |||||||||||||||||||| |||||||| | |:|||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              |||| |||||||||||||||||||||||||| ||  ||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              |||| |||||||||||||||| |||||  |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              |||||  |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||| ||||||||||||||||||||||| |||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              |||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              |||||||||  | |||||||| |||||||||||||||||||||||||| |||||||||  |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              |||||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
```

```
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             |||||||||||||  ||||||||||||||||    ||||||||||||||||||||  |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||||||| |||||||  |||||||||| |
Db       1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||||:|||||||||  |:||||||||||||||||||||||||
Db       1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||||||| |||| |||||||||||||||:|||||
Db       1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:||||||:||||||| |||||| | || ||||||||||
Db       1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1321 CTCAGCCAATGGATGCCCTGGACTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTT 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
               |||||||||||||||||||||||| :|| |||||||||| ||||||||||||||||||
Db       1381 TTACTCCTCTTTGGACCCTGTATCTTCAACTTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
             ||||||||:|| |||             |||||||| ||| ||||||||||||| ||||||
Db       1441 GAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAACCCCAGATGCAGTCCATGACTAAA 1500

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
             |||||||| |||||||||||||||||||:||| | | | |||||||||||||||||  ||
Db       1501 ATCTACCGTGGACCCCTGGACCGGCCTGCTAGACTATGCTCTGATGTTAATGACATTGAA 1560

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
             | |||||||||| |||||||||||| |||||||||| ||||||| | |||||||||  ||||
Db       1561 GTCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTACACTCCAATTCAGTAGGA 1620

Qy       7189 AGCAGTTAG 7197
             |||||||||
Db       1621 AGCAGTTAG 1629


RESULT 131
BD196237
LOCUS       BD196237               1422 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196237
VERSION     BD196237.1  GI:33006007
KEYWORDS    JP 2002512530-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 3 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/3
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
           MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
           Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .1422
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1422
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             |||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||||:|||:|||||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||||:|||||:||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||| :|:|:|||||:|||||||||||||||||||||||||||:||||:|||||
Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             |||||||||||:|||||||||||||:::|||||||||||||||||||||||||||||||||
Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:||||||
Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             ||||||||||:||||:|||||||||||:|||||:|||||:|||||||||||||||||||||
Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||||||||||||||||||:|||||||:::|||||||||||||||||||||||||
Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             |||||||||||||||||:||||||||||:||||||||:|||||||||||||:||:||
Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |::|||||||||||||||:||:||||||:||||||||||||||||:|||:|||||:||||
Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
             |||||||::||||::||||||||:||||||||:||||||||||||||||||||:|:||||
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:|||||||||||||||||||||::|||||||||||||||||||||||||||||||
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
```

```
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807

Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867

Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927

Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987

Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047

Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107

Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167

Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227

Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287

Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337

Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 132
BD267479
LOCUS       BD267479               1422 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267479
VERSION     BD267479.1  GI:33077247
KEYWORDS    JP 2002534980-A/20.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 20 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/20
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
              disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .1422
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1422
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967

Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027

Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087

Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147

Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207

Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267

Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327

Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387

Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447

Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507

Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567

Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627

Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687

Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839
```

```
Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
                ||||||:||:|||||||||||||:|||||||||||||||||||||||||||||||||||
Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
                |||:||||:||||||:|||||||||||:|||||||||||||||||||||||||||||||
Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
                |||||||||||||||||  |||||:|||||||||||||||||||||||||||||:||||
Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
                ||||  |||||||||||||||||||||||||||:||||:|||||||||||||||:|||:|:|
Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
                |||||||||||||||||||:||||||||:|||||||||||||||||||:|:|||||||:|
Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
                |||||||||||||||||||::||||||||:|:|||||||||||||||||||||||:|||:|
Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227
                ||||||||||:|||||||||||||||||||||:|||||||:||||:|||||||||||||:
Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287
                ||||||||:|||||||||||||:|||||||||| |||||||||:|||||||||||:|||:
Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337
                |||:|||||||||:||||||||:|||||||||||||||||||||||||
Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 133
AX000958
LOCUS       AX000958               1422 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 3 from Patent WO9902696.
ACCESSION   AX000958
VERSION     AX000958.1  GI:7241200
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 3 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .1422
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
                ||||||||||:|||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
                |||||||||||||||||||:|||:|||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
                |||||||||||||||||||||||||||:|||||:|||||||||||:|||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
                |||||||| :|:|:|||||:|||||||||||||||||||||||||:||||:|||||
Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
                |||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||
Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
                :|||||||||||||||||||||||:|::|||||||||:|||:|||:|||||||||||||
Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
                ||||||||||:||||:|||||||||||:|||||:|||||:||||||||||||||||||||
Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
                |||||||||||||||:|||||||||:|:||||||:::|||||||||||||||||||||||
Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
                |::|||||||||||||||:||:||||||:|||||||||||||||:|||:||||||:||||
Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
                ||||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||
Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
                ||||||:||||||||||||||||||||||::|||||||||||||||||||||||||||||
Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
                ||:|||||||||||||||||||||||||||||:|||||:||||||||||||||||||||
Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                |||||||||||||||||||:|||||||||||:|||||||||||||||||||:|||||
Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
                |||||||:|||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
                ||||||:||:|||||||||||||:|||||||||||||||||||||||||||||||||||
Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
                |||:||||:||||||:|||||||||||:|||||||||||||||||||||||||||||||
Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
                |||||||||||||||||  |||||:|||||||||||||||||||||||||||||:||||
Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
                ||||  |||||||||||||||||||||||||||:||||:|||||||||||||||:|||:|:|
Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
                |||||||||||||||||||:||||||||:|||||||||||||||||||:|:|||||||:|
```

```
Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167

Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227

Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287

Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337

Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 134
AX027472
LOCUS       AX027472               1422 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 22 from Patent FR2788784.
ACCESSION   AX027472
VERSION     AX027472.1  GI:10188436
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 22 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967

Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027

Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087

Db        121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147

Db        181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207

Db        240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267

Db        300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327

Db        360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387

Db        420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447

Db        480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507

Db        540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567

Db        600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627

Db        660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687

Db        720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807

Db        840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867

Db        898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927

Db        955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987

Db       1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047

Db       1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107

Db       1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167

Db       1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227

Db       1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGR 4287

Db       1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337

Db       1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 135
AF072501
LOCUS       AF072501               1422 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
```

```
ACCESSION   AF072501
VERSION     AF072501.1  GI:4262285
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 1422)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1422
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.7A16"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     misc_feature    1. .1422
                     /note="corresponds to pol"
     repeat_region   1. .1422
                     /rpt_family="HERV-W"

  Query Match             17.7%;  Score 1329;  DB 1;  Length 1422;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches 1306;  Conservative  115;  Mismatches    1;  Indels    8;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAGTTATCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||:||||||||||||||:||:|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTTTTAGCTGCCTGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||:|||||:||||||||||:||||||||||||:||
Db         121 CCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||| :|:|:||||||:|||||||||||||||||||||||||||:||||:|||||
Db         181 TTTCCAAA-GAGAAGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGACTAAGATTAT 239

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||||
Db         240 CCAAAGGCACCAAGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATC 299

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||
Db         300 TCAAAACCCTAAAGCAACTAAGAGAGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 359

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             ||||||||||:||||:|||||||||||:|||||:|||||:||||||||||||||||||||
Db         360 TTCCCCAGGTATGGCAAAATAGCCAGGCCATTATATATACAGTAATTAAGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||||||||:|||||||||:|:||||||:::||||||||||||||||||||||||
Db         420 GCCAATACCCATTTAATAAGATGGATACCTGAAGCCAAAGTGGCTTTCCAGGCCCCTAAA 479

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
             ||||||||||||||||:||||||||||:||||||||:||||||||||||||:||:||
Db         480 GAAGGCCTTAAACCCAAGTCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTAT 539

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |::|||||||||||||||||:||:||||||:||||||||||||||:|||:|||||:||||
Db         540 ACATCACAGAAAAAAACAGAAACAGCTCTGGGAGTCCTTACACAGGTCCAAGGGACGAGC 599

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
             ||||||||::||||::||||||:||||||||:||||||||||||||||||||:|:||||
Db         600 TTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCTTCAT 659

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db         660 TGTTTATGGGTAGTGGTGGCAGTAGCAGTTGTAGTATCTGAAGCAGTTAAAATAATACAG 719

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:|||||||||||||||||||||||||||||:|||||::||||||||||||||||||||
Db         720 GGGAGAGATCTTACTGTGTGGACATCTCATGAGGTGAACAGCATACTCACTGCTAAAGGA 779

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||:||||||||||||:|||||||||||||||||||:|||||
Db         780 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 839

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:|||||||||||||||||||||||||||||:|||||||||||||||||||||
Db         840 GTGCTGCAACTGTGCACTTGTGCAACTCTTAACCCAGTCNCATTTCTTCCAGACAATG-- 897

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
                  ||||||:||:|||||||||||||:|||||||||||||||||||||||||||||||
Db         898 ---AAGATAGAATATAACTGTCAACAAATAATTTCTCAAACCTATGCCACTCGAGGGGAC 954

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||||:|||||||||||||||||||||||||||||||
Db         955 CTTCTAGAAGTTCCCTTGACTGATCCTGACCTTCAACTTGTATACTGATGGAAGTTCCTT 1014

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
             |||||||||||||||| |||||:|||||||||||||||||||||||||||||:||||
Db         1015 TGTAGAAAAAGGACTTC-AAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATATTTGA 1073

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
             |||| ||||||||||||||||||||||||||:||||:|||||||||||||:|||:|:|
Db         1074 AAGT-ATCCCCTCACTCCAGGAACTAGTGCTTAGCTGGCAGAACTAATAGCCTTCATTGG 1132

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
             |||||||||||||||||||:||||||||:|||||||||||||||||||:|:||||||:|
Db         1133 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTGAGTATGCTCA 1192

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
             |||||||||||||||||||||::|||||||:|:|||||||||||||||||||||:|||:|
Db         1193 CCTAGTCNTCCATGCCCATGAGGCAATATGCAGAGAAAGGGAATTCCTAACTTCCGAGGG 1252

Qy       4168 AACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTAR 4227
             ||||||||||:||||||||||||||||||:|||||:||||:||||||||||||||:
Db         1253 AACACCTATCACACATCAGGAAGCCATTAGGAGATTATTACTGGCAGTACAGAAACCTAA 1312

Qy       4228 AGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAAGGGAAATASAAGR 4287
             ||||||||:|||||||||||||:|||||||||| ||||||||:||||||||||:|||:
Db         1313 AGAGGTGGAAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGAAAAGGGAAATAGAAGG 1372

Qy       4288 GAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTC 4337
             |||:|||||||||:|||||||||:|||||||||||||||||||||||||
Db         1373 GAATTGCCAAGCAGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTC 1422


RESULT 136
BD136197
LOCUS       BD136197               1481 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION   BD136197
VERSION     BD136197.1  GI:23231142
KEYWORDS    JP 2002509437-A/7.
SOURCE      unidentified
  ORGANISM  unidentified
```

```
            unclassified sequences.
REFERENCE   1  (bases 1 to 1481)
  AUTHORS   Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
            Perron,H. and Mandrand,B.
  TITLE     Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: JP 2002509437-A 7 26-MAR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002509437-A/7
            PD   26-MAR-2002
            PF   07-JUL-1998 JP 1999508255
            PR   07-JUL-1997 FR     97/08816
            PI   GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC  PI
              BEDIN,
            PI   MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,
            PI   BERNARD MANDRAND
            PC   C12N15/48,C12Q1/70,C07K14/15,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Retroviral nucleic material and nucleotide fragments, in  CC
                 particular,
            CC   associated with multiple sclerosis and/or
            rheumatoid arthritis,
            CC              for
            CC   diagnostic, prophylactic and therapeutic uses FH   Key
                 Location/Qualifiers
            FT   source          1. .1481
            FT                   /organism='Unidentified'.
FEATURES                Location/Qualifiers
     source             1. .1481
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||| ||||||||:||:||||:| |:|||| :|||| ||||||||
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             |||||||||||||| ||||||  |||||||||||||:|||:|||||||||||| |||
Db         61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             | || || ||| |||||||||||||||||||||| ||||||||| | |:||||||| | ||
Db        121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             |||| |||||||||||||||||||||||||| ||  |||||||||||||||||||||| ||
Db        301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             |||| |||||||||||||||| ||||| |||| |||||||||||| |||:||| ||||
Db        361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             |||||||||||||| |||||||||||||||||||||||| |||||||||||||||| |||
Db        481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             |||||||||||||||||||||||| ||||||||||| ||||||:|||||| ||||||||
Db        541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             |||||||||| | |||||||||| ||||||||||||||||||||||||||| ||||||||| |
Db        721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             |||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
             |||||||||||||||||| | ||||| || |||||||| ||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
             |||||||||||||| |||||||||||||||||   |||||||||||||||||||| |||||
Db        961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||| |
Db        1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
             |:||||||||||||||||||||:|||||||||  |:|||||||||||||||||||||||||
Db        1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
             || |||||| |||||||||||||||||||||| |||| ||||||||||||||:||||||
Db        1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
             ||||:|||||||||||||| |||:||||||:||||||| |||||| | || ||||||||||
Db        1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
             |||||||:|||||||||||||| ||||||||||||||||||||||||||||| ||||||||
Db        1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
              |||||||||||||||||||||||||:|||||||||||||| ||||||||||||||||||
Db        1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
             |||||||:|| ||||| ||||      | ||| |  || |
Db        1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481

RESULT 137
CS248641
LOCUS       CS248641               1481 bp    DNA     linear   PAT 09-JAN-2006
DEFINITION  Sequence 9 from Patent EP1612270.
ACCESSION   CS248641
```

```
VERSION     CS248641.1  GI:84662036
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Parahnos-Baccala,G.
  TITLE     Retroviral nucleic material and nucleotide fragments, in particular
            associated with multiple sclerosis and/or rheumatoid arthritis, for
            diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: EP 1612270-A 9 04-JAN-2006;
            BIOMERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1481
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             17.7%;  Score 1324.2;  DB 1;  Length 1481;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches 1365;  Conservative   24;  Mismatches   92;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
Db          1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCCTTTCGCTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
Db            61 GCACCCCCTCCATGCTGCTGTACAACCAGTAGCTCCCCTTACCAAGAGTTTCTATGAAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
Db           121 ACGCGGCTTCCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db           181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db           241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db           301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db           361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db           421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db           481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db           541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db           601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db           661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db           721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db           781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db           841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db           901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db           961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db          1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db          1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db          1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db          1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900
Db          1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db          1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

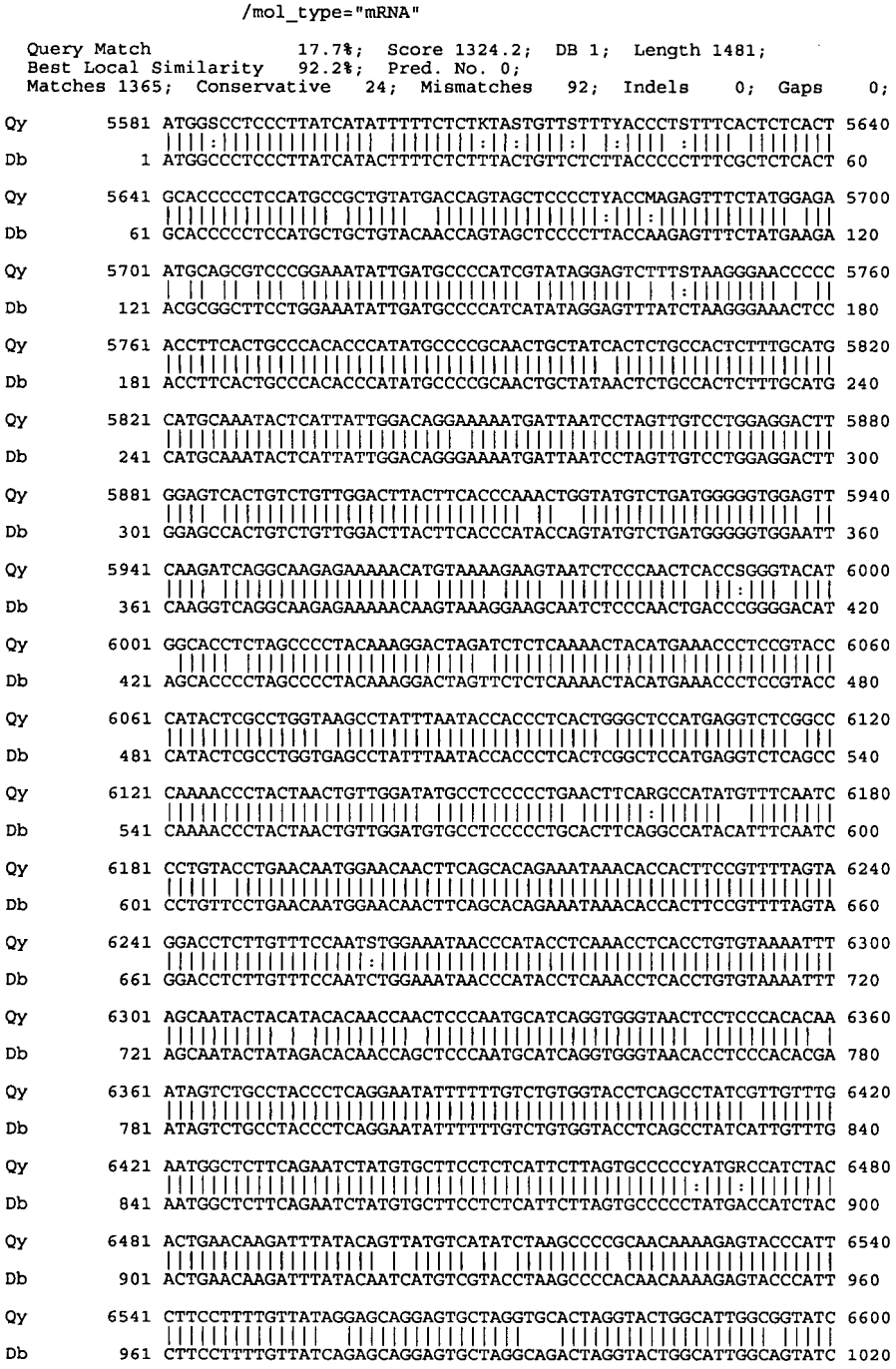
Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db          1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

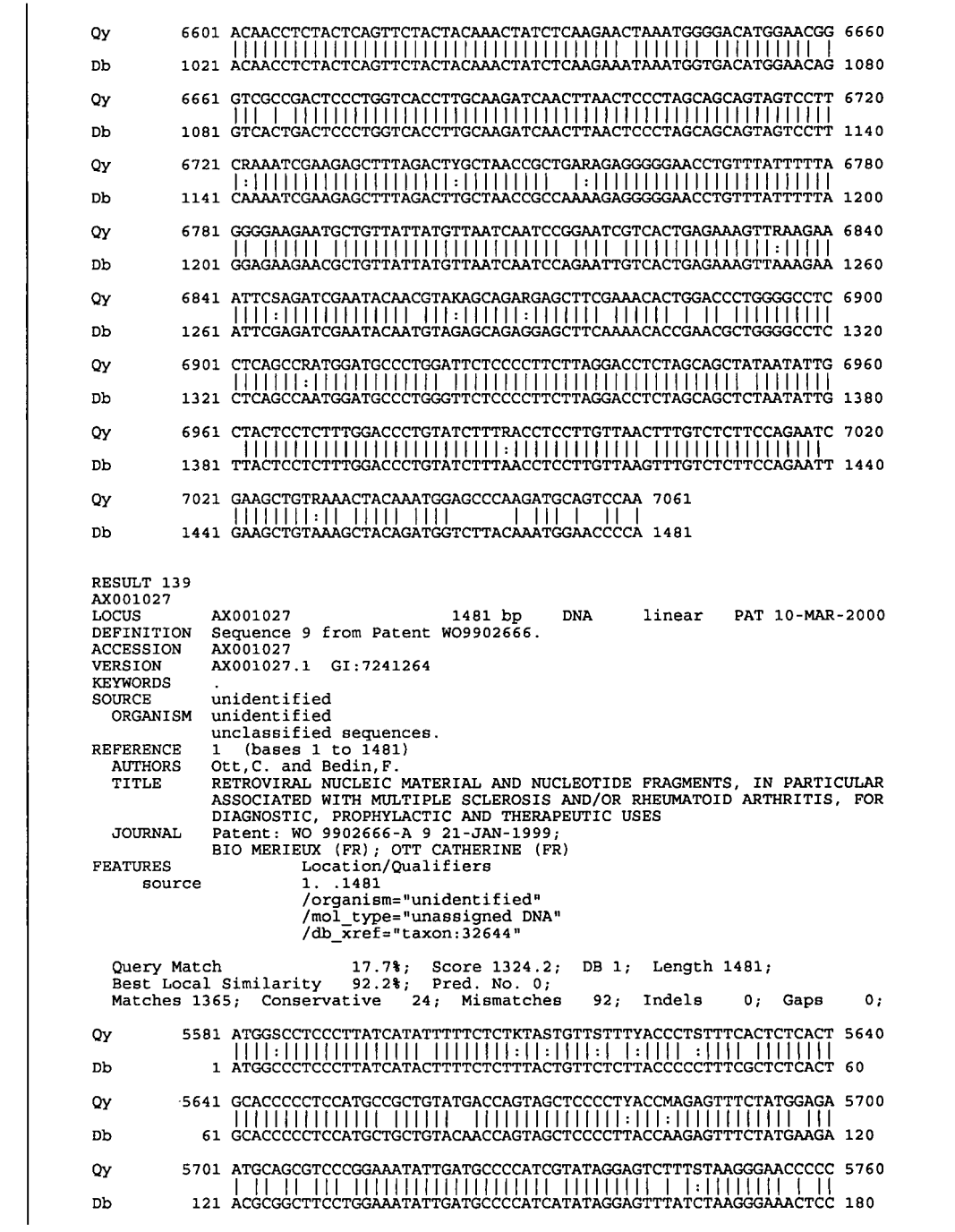
Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
Db          1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481


RESULT 138
AR344387
LOCUS       AR344387               1481 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 105 from patent US 6582703.
ACCESSION   AR344387
VERSION     AR344387.1  GI:33740328
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1481)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 105 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1481
                     /organism="unknown"
```

/mol_type="mRNA"

Query Match 17.7%; Score 1324.2; DB 1; Length 1481;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;





RESULT 139
AX001027
LOCUS AX001027 1481 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 9 from Patent WO9902666.
ACCESSION AX001027
VERSION AX001027.1 GI:7241264
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 1481)
AUTHORS Ott,C. and Bedin,F.
TITLE RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
JOURNAL Patent: WO 9902666-A 9 21-JAN-1999;
BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES Location/Qualifiers
source 1. .1481
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 17.7%; Score 1324.2; DB 1; Length 1481;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;

```
Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
Db         241 CATGCAAATACTCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
Db         301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAGTATGTCTGATGGGGGTGGAATT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
Db         361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
Db         421 AGCACCCCTAGCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
Db         481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
Db         541 CAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAATC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
Db         601 CCTGTTCCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
Db         661 GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
Db         721 AGCAATACTATAGACACAACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
Db         781 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
Db         841 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540
Db         901 ACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATT 960

Qy       6541 CTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 6600
Db         961 CTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020

Qy       6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660
Db        1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

Qy       6661 GTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db        1081 GTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Qy       6721 CRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTA 6780
Db        1141 CAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAACCTGTTTATTTTTA 1200

Qy       6781 GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA 6840
Db        1201 GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA 1260

Qy       6841 ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTGGACCCTGGGGCCTC 6900

Db        1261 ATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320

Qy       6901 CTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960
Db        1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

Qy       6961 CTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATC 7020
Db        1381 TTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATT 1440

Qy       7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAA 7061
Db        1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
```

RESULT 140
AR344431
LOCUS AR344431 1600 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 207 from patent US 6582703.
ACCESSION AR344431
VERSION AR344431.1 GI:33740372
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G., Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
TITLE Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting
JOURNAL Patent: US 6582703-A 207 24-JUN-2003; Bio Merieux; Marcy l'Etoile; FRX;
FEATURES Location/Qualifiers
source 1. .1600
/organism="unknown"
/mol_type="genomic DNA"

Query Match 17.6%; Score 1320.4; DB 1; Length 1600;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1402; Conservative 111; Mismatches 87; Indels 16; Gaps 9;

```
Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
Db            61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
Db           120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
Db           179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
Db           239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
Db           299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
Db           352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
```

```
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
              ||||||||:|||||||||||:|||||||||| ||||||||||||||| ||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
              |||||||||||||||||||||| |||  ||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589

Qy       2896 TACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 2955
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        590 TACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGA 649

Qy       2956 GYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYG 3015
              |||||||:||||||||||||:||||||:|:|||||| :|||||||||||||||||:|:::
Db        650 GYCARTYYTCATACCTGGACAYTCTTGTYCYTCRGTACRTGGATGATTTACTTTTAGYCR 709

Qy       3016 CCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCT 3075
              ||:|||||||||||||||||||||||||||||||::||||||:|:||||||::|:||||
Db        710 CCCRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGMACTCTTAAMTTTCCTYRCTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135
              |||||||||:||||||||||||:|:|:|||||::|||||||||||||:||:|||:|||||||
Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCRGCTCTGCTCACAGSAGRTTARATACTTAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
              |:||||:|||||||||||||||||:::|||||||||||||:::||||||||||||||||
Db        830 GGCTAAAATTATCCAAAGGCACCAGRRCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255
              :||||||||||:|||||||||||||||||::|::|||||||:|||:|||:||||
Db        890 STTATCCTCATCCCAAAACCCTAAAGCAACTAASAGSGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315
              |||:||||::|||| |||||||:  :|::::|||||| :|||||||||||::|||||:|
Db        950 GCCRAAWAYRGATT-CCCAGGTACASCMMRRTAGCCAGACCATTAWATACACKAATTARG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375
              ||||||||:|||||||:||||:|||||:|||||||||:|:|:||||||:::||||||||||
Db       1009 GAAACTCARAAAGCCARTACCYATTTAGTAAGATGGACAYCTGAAGCAGAAGTGGCTTTC 1068

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435
              |||||||  ||  | | |||  ||||||||:|||||||| ||:|||||||:|||||||
Db       1069 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1126

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495
              | ||||:||:|||:: ||||| |||||||||:||:||||||:|||||||||| |||:||
Db       1127 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1185

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555
                 |||||:|||||||||||::||| ::||||||:|||||||:|||||||||||||||||
Db       1186 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1245

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
              |||||:|:|||| |||||:||||| ||| ||||||||||||:: ||||||||||||||||
Db       1246 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1305

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
              |||||||||||||||:|||||||| ||||||||||||||||||:|||||::|||||||||
Db       1306 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1365

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
              |||||||||||||||||||| ||||||||||: |||||||| |:|||||||||||||||
Db       1366 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1425

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795
              ||||||: |||||||||| :|||| ||||||||||||||||| || |:| ||||||||||
Db       1426 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1485

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855
              |||||||||||||||||||:||:|||||||||||||:||||| ||||||||||||||
Db       1486 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1545
```

```
Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
              ||||||||||||||:||||:|||||:||||||||||:|||| ||||||||||||
Db       1546 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1600


RESULT 141
AR344430
LOCUS       AR344430               1600 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 206 from patent US 6582703.
ACCESSION   AR344430
VERSION     AR344430.1  GI:33740371
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1600)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 206 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1600
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             17.5%;  Score 1309.8;  DB 1;  Length 1600;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches 1383;  Conservative  128;  Mismatches   89;  Indels   16;  Gaps    9;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
              ||||||| |||||||||||||| |||||||||||||||||| |||||||||||||||||
Db          1 ATTATGCCTGAAAGCCCCACTCCCTTGTTAGGGAGAGACATTTTAGCAAAAGCAGGGGCC 60

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
              |||||||||||||||||||| |||||| |||| |||| ||| ||||||||||||||||||
Db         61 ATTATACACCTGAACATAGGAAAAGGAATACCCATTTGCTGT-CCCCTGCTTGAGGAAGG 119

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        120 AATTAATCCTGAAGTCTGGGCAATAGAAGGACAATATGGACAAG-CAAAGAATGCCCGTC 178

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
              |||||||||||||||||||||||| | |||||||||||||||||| ||||||| || | |
Db        179 CTGTTCAAGTTAAACTAAAGGATTCTGCCTCCTTTCCCTACCAAAGGAAGTACCCTCTTA 238

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAG 2595
              ||||| |||||| ||||||| ||||||||||||||||||| |||||||||||||| |||
Db        239 GACCCGAGGCCCTACAAGGANCTCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAG 298

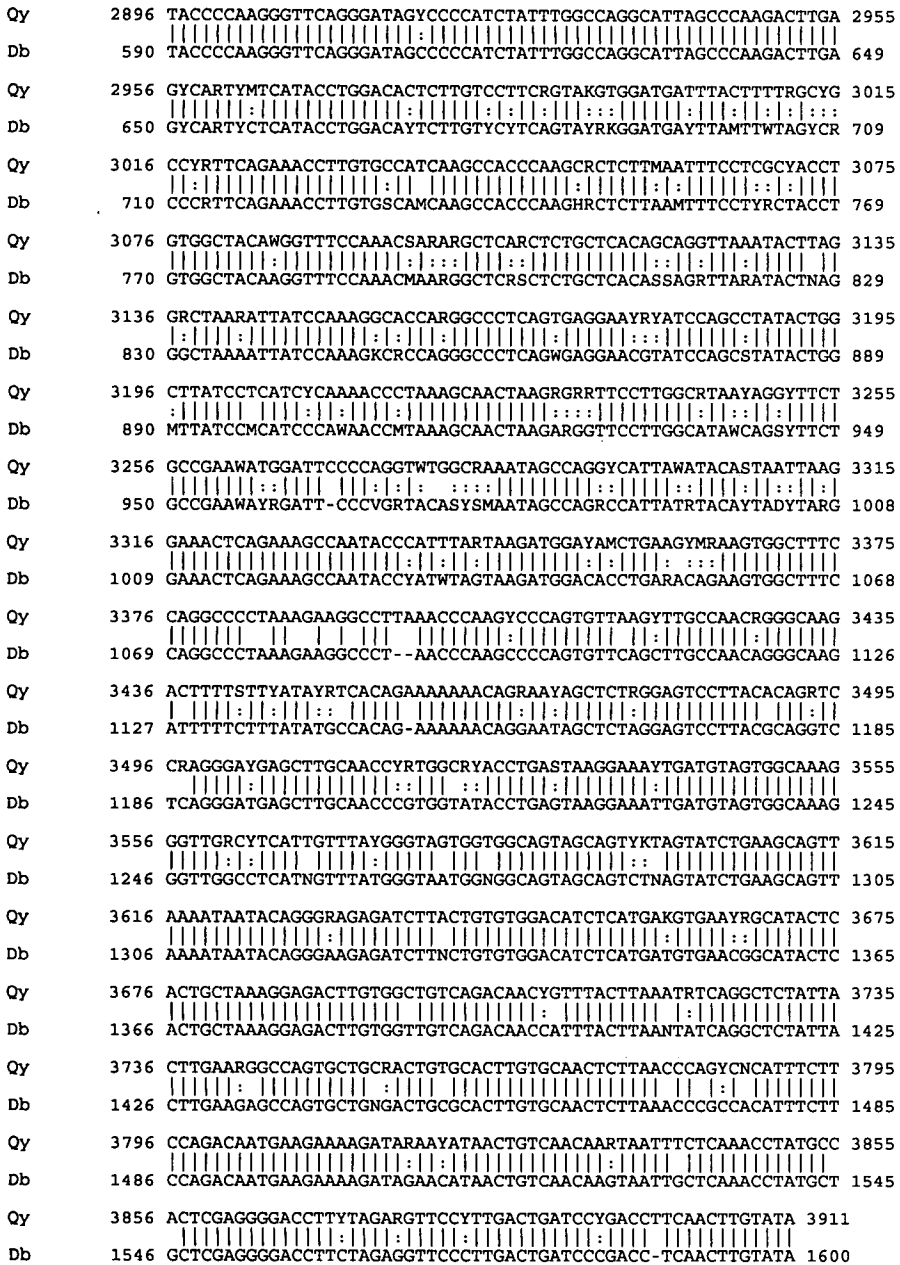
Qy       2596 TAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCC 2655
              |||||||||||| | ||||||| || | |         ||| | | | |||||||
Db        299 TAAAACCATGCAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTAAGGAAACCC 351

Qy       2656 AGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTT 2715
              | |||||||||| |||||||||||||| |||||||||||||||| ||||| ||||| | |
Db        352 AACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTATCAAT-GAGGCTGTTGTTCCTC 409

Qy       2716 TATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGT 2775
              |||||||||||||||||| ||||||||| |||::|||||||||||||||||||||||||||
Db        410 TATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCCAAATACCAGAGGAAGCAGAGT 469

Qy       2776 GGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTC 2835
              |||||||:|||||||||||:|||||||||| ||||||||||||||| ||||||||||||
Db        470 GGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTC 529

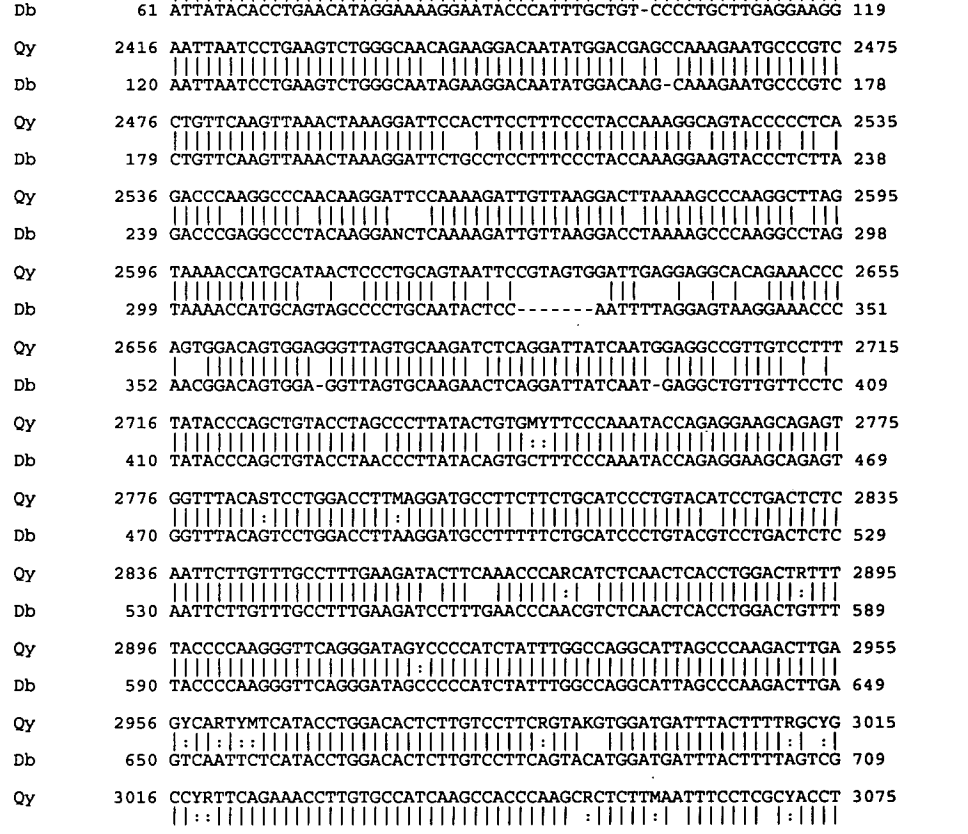
Qy       2836 AATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTT 2895
              |||||||||||||||||||| |||  |||||||:| ||||||||||||||||||:|||
Db        530 AATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTT 589
```



RESULT 142
AR344429
LOCUS AR344429 1597 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 205 from patent US 6582703.

ACCESSION AR344429
VERSION AR344429.1 GI:33740370
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1597)
AUTHORS Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G., Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
TITLE Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting
JOURNAL Patent: US 6582703-A 205 24-JUN-2003;
Bio Merieux; Marcy l'Etoile;
FRX;
FEATURES Location/Qualifiers
source 1. .1597
/organism="unknown"
/mol_type="genomic DNA"

Query Match 17.2%; Score 1291.9; DB 1; Length 1597;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 1405; Conservative 88; Mismatches 104; Indels 19; Gaps 10;



```
Db        710 CCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCT 769

Qy       3076 GTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAG 3135

Db        770 GTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAG 829

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195

Db        830 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGG 889

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCT 3255

Db        890 CTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCT 949

Qy       3256 GCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAG 3315

Db        950 GCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANG 1008

Qy       3316 GAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTC 3375

Db       1009 GAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTC 1065

Qy       3376 CAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAG 3435

Db       1066 CAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAG 1123

Qy       3436 ACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTC 3495

Db       1124 ATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTC 1182

Qy       3496 CRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAG 3555

Db       1183 TCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAG 1242

Qy       3556 GGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615

Db       1243 GGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTT 1302

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675

Db       1303 AAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTC 1362

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735

Db       1363 ACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTA 1422

Qy       3736 CTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTT 3795

Db       1423 CTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACTCTTAAACCCGCCACATTTCTT 1482

Qy       3796 CCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCC 3855

Db       1483 CCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCT 1542

Qy       3856 ACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911

Db       1543 GCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATA 1597


RESULT 143
BD196235
LOCUS       BD196235               1321 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196235
VERSION     BD196235.1  GI:33006005
KEYWORDS    JP 2002512530-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1321)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 1 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/1
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC           diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .1321
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360

Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420

Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480

Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540

Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600

Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 660

Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659
```

```
Qy       661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
            |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy       721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy       781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy       841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy       901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
            ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db       900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy       961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy      1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db      1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy      1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
            |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy      1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
            |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db      1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy      1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy      1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
            ||||||||:|||||||||||||||||||||||||||||||||||     ||||||||||||
Db      1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy      1321 GAAG 1324
            |||| |
Db      1317 GAGG 1320
```

```
RESULT 144
BD267477
LOCUS       BD267477               1321 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267477
VERSION     BD267477.1  GI:33077245
KEYWORDS    JP 2002534980-A/18.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1321)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 18 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/18
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .1321
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

Query Match 17.0%; Score 1270.4; DB 1; Length 1321;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1291; Conservative 22; Mismatches 7; Indels 4; Gaps 3;

```
Qy         1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy        61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy       121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy       181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy       241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db       241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy       301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
             |||||| |||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy       361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db       361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy       421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
             |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db       421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy       481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
             ||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy       541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
             |||||||||||||||||||||||||||||||:|||||||||:||||:::|||||||||||
Db       540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy       601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy       661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
             |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy       721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db       720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779
```

```
Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:|||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              || |
Db       1317 GAGG 1320
```

```
RESULT 145
AX000956
LOCUS       AX000956               1321 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 1 from Patent WO9902696.
ACCESSION   AX000956
VERSION     AX000956.1  GI:7241198
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1321)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 1 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              |||||||||||||||||||||||||||||||:||||||||:|||:::||||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198
```

```
Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
             ||||||||:||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
             || |
Db       1317 GAGG 1320


RESULT 146
AX027470
LOCUS       AX027470               1321 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 20 from Patent FR2788784.
ACCESSION   AX027470
VERSION     AX027470.1  GI:10188434
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 20 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              ||||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||||||::||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              || |
Db       1317 GAGG 1320


RESULT 147
AF072500
LOCUS       AF072500               1321 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W sequence.
ACCESSION   AF072500
VERSION     AF072500.1  GI:4262284
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1321)
```

```
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 1321)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.6A2"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     misc_feature    1. .1321
                     /note="corresponds to gag"
     repeat_region   1. .1321
                     /rpt_family="HERV-W"
     5'UTR           1. .120
                     /note="5' end of R uncertain"
     5'UTR           121. .574
                     /note="U5"
     primer_bind     578. .595
                     /note="putative primer binding site for tRNA-W"

  Query Match             17.0%;  Score 1270.4;  DB 1;  Length 1321;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches 1291;  Conservative   22;  Mismatches    7;  Indels    4;  Gaps    3;

Qy          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCC 60

Qy         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120

Qy        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180

Qy        181 TGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTC 240

Qy        241 CGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        241 CGCTGTGCTCCTGATCCAGCGAAGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGC 300

Qy        301 CATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT 360
              |||||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        301 CATTGTTCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACT 360

Qy        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA 420

Qy        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        421 TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATAC 479

Qy        481 GARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGGAAGTGGTTCACCACCATC 540
              ||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        480 GAAGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTGGTTCACCACCATC 539

Qy        541 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTRACATTTTGGCRACCAMSRACGGACATCC 600
              ||||||||||||||||||||||||||||||:||||||||||:||||:::||||||||||
```

```
Db        540 TTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAACGGACATCC 599

Qy        601 MAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGARAAT 660
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        600 AAAGTGATGGGAAACGTTCCCCGCAAGACAAAAACGCCCCTAAGACGTATTCTGGAAAAT 659

Qy        661 TGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGAACAATTTGACCCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 719

Qy        721 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        720 GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTTG 779

Qy        781 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAA 839

Qy        841 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC 899

Qy        901 CTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAATGG 960
              ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db        900 CTATCCCAGCATCCCCGACTCCTTCCCCACTTAATAAGGACCCCCCTTCAACCCAAATGG 959

Qy        961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAAT 1019

Qy       1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYT 1080
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||:|
Db       1020 TATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCCT 1078

Qy       1081 TTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCT 1140
              |||::|||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1079 TTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAACCCT 1138

Qy       1141 GATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1200
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GATGGCTATATTGGTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAGAT 1198

Qy       1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACTGC 1258

Qy       1261 AGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA 1320
              ||||||||:|||||||||||||||||||||||||||||||||||     ||||||||||||
Db       1259 AGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACA 1316

Qy       1321 GAAG 1324
              || |
Db       1317 GAGG 1320
```

```
RESULT 148
CQ073417
LOCUS       CQ073417                1472 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 9217 from Patent WO0157278.
ACCESSION   CQ073417
VERSION     CQ073417.1  GI:41043286
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 9217 09-AUG-2001;
```

```
          Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1472
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN HELA, SIGNAL = 1.3"

  Query Match             16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||||  | ||||  ||||||||||||:||||:||||||||||||||||:
Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              ||||||||||||||||:::||||||||  ||||||||||||||||||||:||||||||||||
Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::||| |||||:|| :|||:||||||||||:|||||| |||||||: |
Db         121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:|||||||  |||:|||||:|||||:|||||  ||||||||||||||||||||||||||||
Db         180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:||||||||:|:||||||:::||| ||||||||||||||  ||  | | |||  |||
Db         240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              |||||:||||||||   ||||:||||||||:|||||||||||| :||:|||::||||||
Db         298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||||||||:||:||||||:||||||||||||||||:|||:|||||: ||| |||||||::
Db         357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::||||||:||||||||:||||||||||||||||||||:|:||||||||||: ||
Db         417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
              |||||| ||||||||||::||||||||||||||||||||| |||||||||| :|||||||
Db         477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
              |||||||||||||||||||:|||||::||||||||||||||||||||||||||||||||
Db         537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
              ||||||||||:||||||||||||:|||||||||| | ||||||:| ||||||||||:||
Db         597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
              |||||||||||||||||||||||||||:| ||||||||||| ||||||||||||||||
Db         657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
              ||:|||||||||||||:|||||  |||||||| ||||||| ||||||||||:||||:|
Db         717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
              ||||: |||||||||:|||| ||||||||||||||||||||||||| |||||||||
Db         777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
              || |||| |||||:|||||||||||| ||||||||||||||||:|||||||||||||||
Db         836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
              ||||||||||||||||||||||:||||:|||||||||||||||:|||:|  |||| |||||
Db          896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
              ||   |||||:||||||||:||||||||||||||||||||:|:||||||:|||||||| ||
Db          956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
              ||||||||||::||||||||:|:|||||||| ||||||||||||:|||:||||||||||||
Db         1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
              |:|||||||||||||||||||:||||||||:|||| ||||||||||||:|||||||||:|
Db         1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
              ||||||||||||:|||||||||| |||||||||:||||||||||:|||:|||:||||||
Db         1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
              || :| ||||||||:|||||||||||||||||||||||||||||||||||||||||||  ||
Db         1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
              |||  ||||||  |||||||||  || ||||||||||||||||||||||||  || ||||
Db         1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
              | ||||||||||||||| ||||    ||||| |||||||| ||| || |||||||  ||
Db         1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

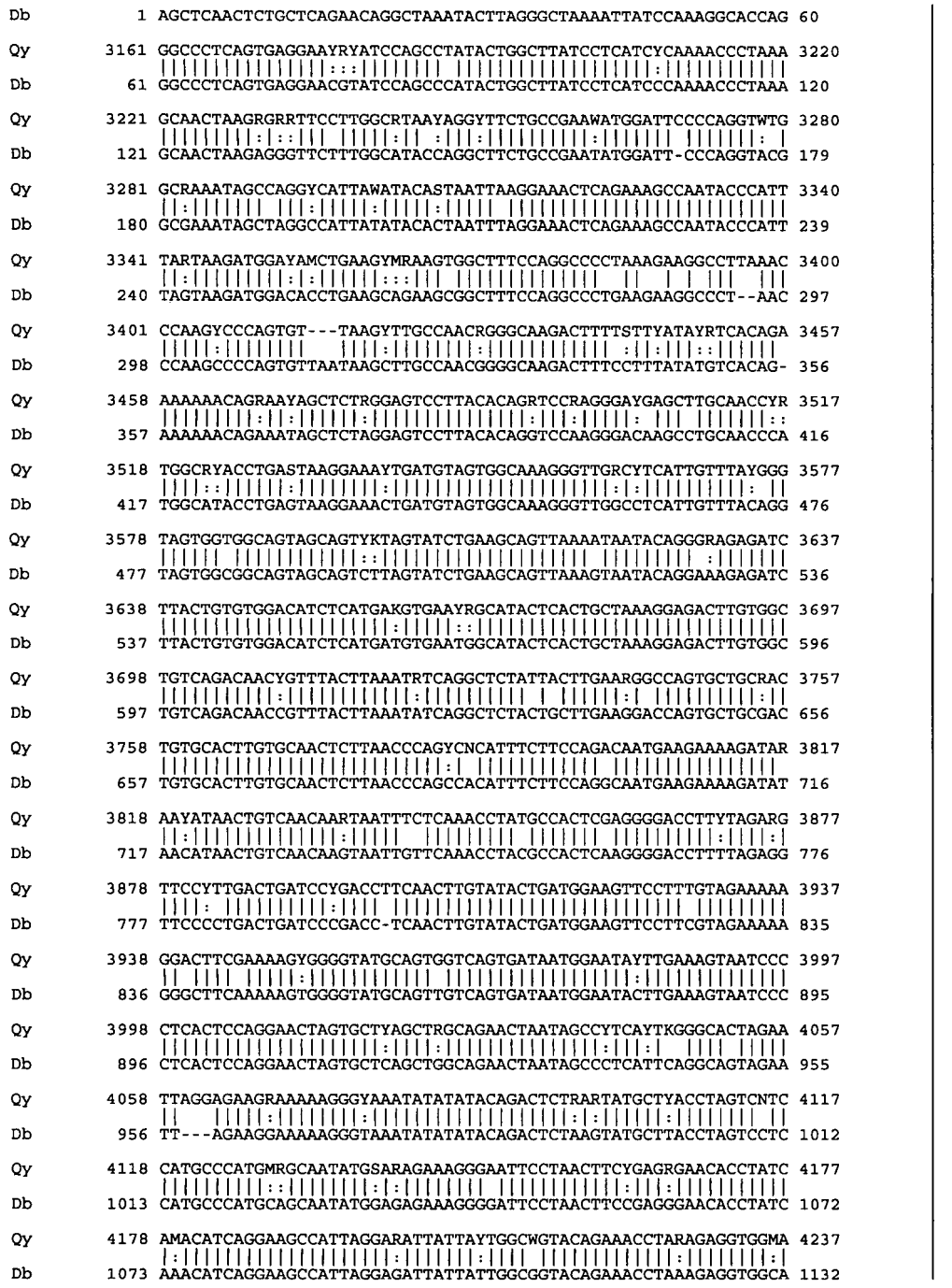
Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
              ||||| |||||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db         1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

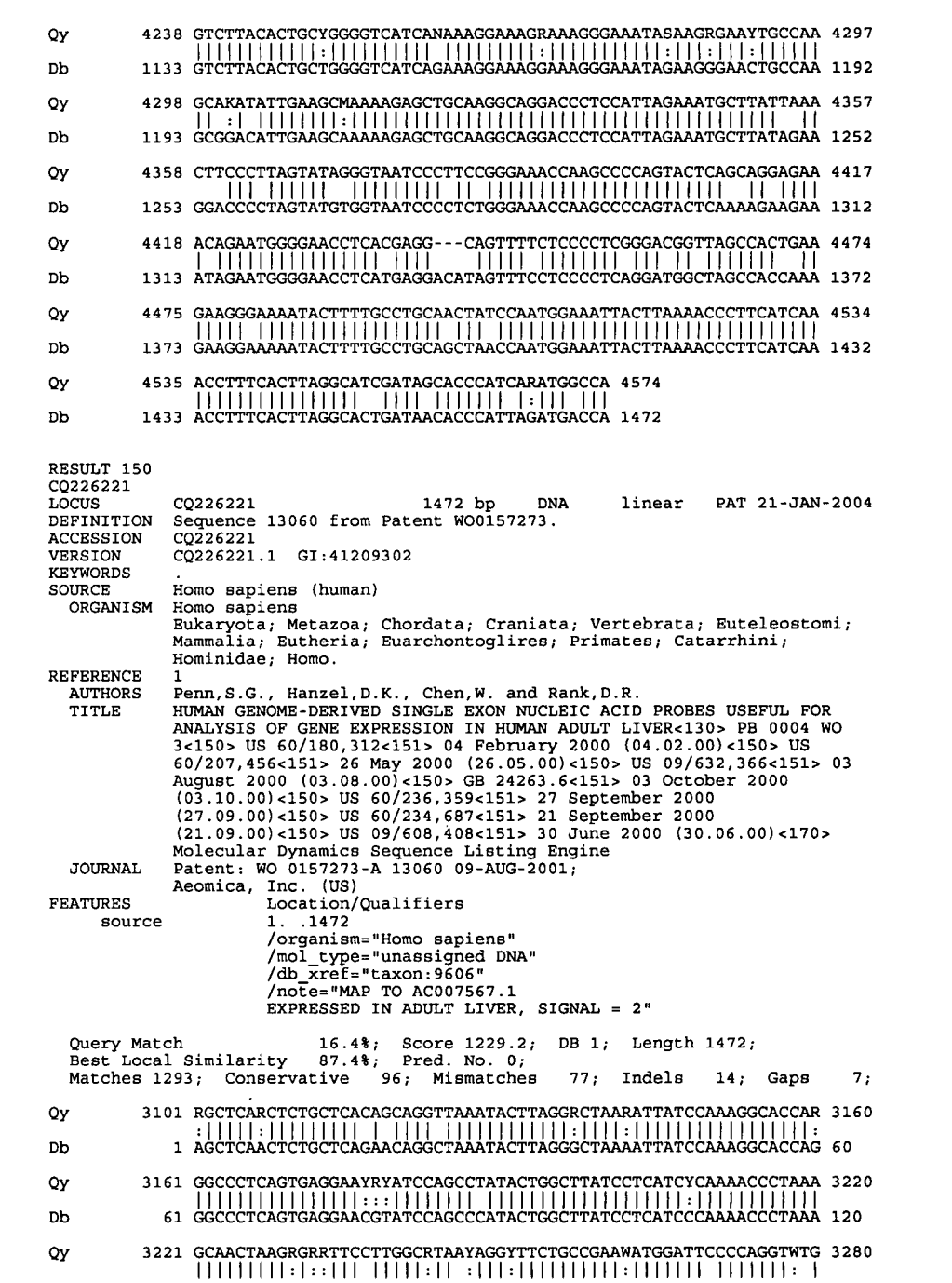
Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
              ||||||||||||||||  |||| |||||||  |:||| |||
Db         1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472

RESULT 149
CQ104281
LOCUS       CQ104281               1472 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 13140 from Patent WO0157272.
ACCESSION   CQ104281
VERSION     CQ104281.1  GI:41073332
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 13140 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1472
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 6"

  Query Match             16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||||  | ||||  ||||||||||||:||||:||||||||||||||||:
```





RESULT 150
CQ226221
LOCUS CQ226221 1472 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 13060 from Patent WO0157273.
ACCESSION CQ226221
VERSION CQ226221.1 GI:41209302
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000 (03.10.00)<150> US 60/236,359<151> 27 September 2000 (27.09.00)<150> US 60/234,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170> Molecular Dynamics Sequence Listing Engine
JOURNAL Patent: WO 0157273-A 13060 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1472
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC007567.1
EXPRESSED IN ADULT LIVER, SIGNAL = 2"

Query Match 16.4%; Score 1229.2; DB 1; Length 1472;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 1293; Conservative 96; Mismatches 77; Indels 14; Gaps 7;

```
Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:||||||| |||:|||||:|||||:||||| |||||||||||||||||||||||||||
Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:||||||||:|:||||||:::||| |||||||||||||   ||  | | |||  |||
Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              |||||:|||||||    ||||:||||||||:|||||||||||| :||:|||::||||||
Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||||||||:||:|||||||:||||||||||||||||:|||:|||||: ||| |||||||::
Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::|||||||:||||||:||||||||||||||||||||:|:||||||||||: ||
Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
              |||||| ||||||||||||::||||||||||||||||||| |||||||||  :|||||||
Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
              |||||||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
              |||||||||||:|||||||||||:|||||||||| | ||||||:| ||||||||||:||
Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
              |||||||||||||||||||||||||:| |||||||||||| ||||||||||||||||
Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
              ||:||||||||||||||:|||||  ||||||||| ||||||| |||||||||||:||||:|
Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
              ||||: |||||||||:|||| |||||||||||||||||||||||||||||||| |||||||||
Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
              || |||| |||||:|||||||||||| ||||||||||||||||:|||||||||||||||
Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
              |||||||||||||||||||||:||||:||||||||||||||:|||:|  |||| |||||
Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
              ||   |||||:|||||||:|||||||||||||||||||:|:|||||:|||||||| ||
Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
              ||||||||||::|||||||:|:|||||||| ||||||||||||:|||:|||||||||||
Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
              |:|||||||||||||||||||:|||||||:|||| |||||||||||||:||||||||:|
Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
              |||||||||||:|||||||||| |||||||||:|||||||||||:|||:|||:||||||
Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
              || :| |||||||:|||||||||||||||||||||||||||||||||||||||||||  ||
Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
              |||  ||||||  |||||||| || |||||||||||||||||||||||  || ||||
Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
              | |||||||||||||||| ||||    ||||| |||||||| ||| || |||||||  ||
Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
              ||||| |||||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
              |||||||||||||||  |||| ||||||| |:||| |||
Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472


RESULT 151
CQ264339
LOCUS       CQ264339               1472 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 12600 from Patent WO0157277.
ACCESSION   CQ264339
VERSION     CQ264339.1  GI:41236871
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 12600 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1472
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 2.3"

  Query Match           16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:||||||||| | |||| ||||||||||||:||||:|||||||||||||||||:
Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              |||||||||||||||:::|||||||| ||||||||||||||||||:||||||||||||
Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::||| |||||:|| :|||:||||||||||:||||||| |||||||: |
Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:||||||| |||:|||||:|||||:||||| |||||||||||||||||||||||||||
Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400
              ||:||||||||:|:||||||:::||| |||||||||||||   ||  | | |||  |||
Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              |||||:|||||||    ||||:||||||||:|||||||||||| :||:|||::||||||
Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
```

```
Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577

Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637

Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697

Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757

Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817

Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877

Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937

Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997

Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057

Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117

Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177

Db       1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237

Db       1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297

Db       1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357

Db       1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417

Db       1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474

Db       1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534

Db       1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574

Db       1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472
```

```
RESULT 152
CQ301436
LOCUS       CQ301436               1472 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 12541 from Patent WO0186003.
ACCESSION   CQ301436
VERSION     CQ301436.1  GI:41262013
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 12541 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1472
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN LUNG, SIGNAL = 1.4"

  Query Match             16.4%;  Score 1229.2;  DB 1;  Length 1472;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 1293;  Conservative   96;  Mismatches   77;  Indels   14;  Gaps    7;

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160

Db          1 AGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 60

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220

Db         61 GGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAAAACCCTAAA 120

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280

Db        121 GCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-CCCAGGTACG 179

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340

Db        180 GCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCAATACCCATT 239

Qy       3341 TARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAAC 3400

Db        240 TAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGGCCCT--AAC 297

Qy       3401 CCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457

Db        298 CCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATATGTCACAG- 356

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517

Db        357 AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCCTGCAACCCA 416

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577

Db        417 TGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAGG 476

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637

Db        477 TAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGATC 536

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697

Db        537 TTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGC 596
```

```
Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
Db        597 TGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGAC 656

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
Db        657 TGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATAT 716

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
Db        717 AACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAGG 776

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAA 3937
Db        777 TTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAAA 835

Qy       3938 GGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCC 3997
Db        836 GGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCCC 895

Qy       3998 CTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAA 4057
Db        896 CTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGAA 955

Qy       4058 TTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTC 4117
Db        956 TT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCTC 1012

Qy       4118 CATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATC 4177
Db        1013 CATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTATC 1072

Qy       4178 AMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMA 4237
Db        1073 AAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGCA 1132

Qy       4238 GTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAA 4297
Db        1133 GTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCAA 1192

Qy       4298 GCAKATATTGAAGCMAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAA 4357
Db        1193 GCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGAA 1252

Qy       4358 CTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAA 4417
Db        1253 GGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGAA 1312

Qy       4418 ACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAA 4474
Db        1313 ATAGAATGGGGAACCTCATGAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACCAAA 1372

Qy       4475 GAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAA 4534
Db        1373 GAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCATCAA 1432

Qy       4535 ACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCA 4574
Db        1433 ACCTTTCACTTAGGCACTGATAACACCCATTAGATGACCA 1472
```

RESULT 153
BD221807
LOCUS BD221807 1326 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses.
ACCESSION BD221807
VERSION BD221807.1 GI:33031577
KEYWORDS JP 2002518051-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1326)
AUTHORS Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A 2 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/2
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC A61P25/00,
PC A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC Nucleic sequence and deduced protein sequence family with CC human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1. .1326
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .1326
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 16.0%; Score 1201.4; DB 1; Length 1326;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1295; Conservative 16; Mismatches 14; Indels 12; Gaps 9;

```
Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777
Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837
Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897
Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
```

```
                ||||  |||||||| ||||||||||||||||||||:|||||||||||||||:|||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
                |||||   ||| | |     ||||| ||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711

Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
                |||||:|||||||||||||||||:|||| ||:||||:|||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
                |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
                |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009

Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
                |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
                |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326


RESULT 154
AR699756
LOCUS       AR699756               1326 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 2 from patent US 6919438.
ACCESSION   AR699756
VERSION     AR699756.1  GI:75205545
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1326)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 2 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1326
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             16.0%;  Score 1201.4;  DB 1;  Length 1326;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1295;  Conservative   16;  Mismatches   14;  Indels   12;  Gaps    9;

Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
              ||||||||||||||||||||||||||||||||||::||||||||||||||||||||||||
Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
              |:||||::|||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
              ||||||||||:|||||||||||||||||||||||||||||||||     ||||||||
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
              |||| |||||||| ||||||||||||||||||:|||||||||||||||:|||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
              |||||   ||| | |     ||||| ||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711

Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
              |||||:|||||||||||||||||:|||| ||:||||:|||||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
              |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009
```

```
Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
             |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
             |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326


RESULT 155
AX007979
LOCUS       AX007979               1326 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 2 from Patent WO9967395.
ACCESSION   AX007979
VERSION     AX007979.1  GI:9995676
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 2 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .1326
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             16.0%;  Score 1201.4;  DB 1;  Length 1326;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1295;  Conservative   16;  Mismatches   14;  Indels   12;  Gaps    9;

Qy        718 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          1 GCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTT 60

Qy        778 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGA 120

Qy        838 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACC 180

Qy        898 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
              ||||||||||||||||||||||||||||||::|||||||||||||||||||||||||||||
Db        181 TCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAA 240

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 300

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 359

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
              |:||||::|||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        360 CCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAC 419

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 479

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 538

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 TGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACA 596

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCA 1377
              |||| |||||||| ||||||||||||||||||||||:|||||||||||||||:|||||||
Db        597 ACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAGACCCTCA 655

Qy       1378 TTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 1437
              |||||   ||| | |      ||||| ||||||||||||||||||||||||||||||||||
Db        656 TTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTAACTTGTG 711

Qy       1438 TGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACC 1497
              |||||:|||||||||||||||||:|||| ||:|||||:||||||||||||||||||||||||
Db        712 TGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGATGTCCACC 770

Qy       1498 ATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 1557
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 ATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATT 829

Qy       1558 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 GAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAA 889

Qy       1618 GTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCC 1676
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        890 GTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTCTGCCGTAGGCCC 949

Qy       1677 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCA 1736
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        950 GGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCA 1009

Qy       1737 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGA 1069

Qy       1797 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGC 1129

Qy       1857 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 CTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTA 1189

Qy       1917 GAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1976
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GAAGTAAGCCGCCCCCTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT 1249

Qy       1977 GCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 2036
              |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||
Db       1250 GCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGG 1309

Qy       2037 ACTGAGGGTGCCTGGGG 2053
              |||||||||||||||||
Db       1310 ACTGAGGGTGCCTGGGG 1326
```

```
RESULT 156
BD136207
LOCUS       BD136207               1511 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION   BD136207
VERSION     BD136207.1  GI:23231152
KEYWORDS    JP 2002509437-A/17.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1511)
  AUTHORS   Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
            Perron,H. and Mandrand,B.
  TITLE     Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: JP 2002509437-A 17 26-MAR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002509437-A/17
            PD   26-MAR-2002
            PF   07-JUL-1998 JP 1999508255
            PR   07-JUL-1997 FR     97/08816
            PI   GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC  PI
              BEDIN,
            PI   MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,
            PI   BERNARD MANDRAND
            PC   C12N15/48,C12Q1/70,C07K14/15,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Retroviral nucleic material and nucleotide fragments, in  CC
                 particular,
            CC   associated with multiple sclerosis and/or
            rheumatoid arthritis,
            CC              for
            CC   diagnostic, prophylactic and therapeutic uses FH   Key
                Location/Qualifiers
            FT   source          1. .1511
            FT                   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .1511
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | ||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||| || |||||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              ||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||| ||||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |   |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      |||||||||||||:|||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || |||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:|||||| |||||||||:|||| ||: |||:|||||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| |||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| ||||||||||||||  |||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db        1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||| ||||||||       | ||||  |||||||
Db        1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||  |||||||||||||||| |||||||||||||||||||||||
Db        1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db        1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||  |||||||||||||||||||||||
Db        1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
```

```
                |||||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
                |||||||||||||||| ||||| ||||||| ||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
                |||||| ||||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 157
BD267460
LOCUS       BD267460               1511 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267460
VERSION     BD267460.1  GI:33077228
KEYWORDS    JP 2002534980-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1511)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 1 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/1
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .1511
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1511
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
                ||||  |||||||||||:|||||||:|||| |||| |||||||  ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
                | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
                |||||||||||||||||||||||||| ||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                  | |||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
                |||||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
                |||||     | |||||||||||||||| ||::|||||||||||||||||| |||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
                |||||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
                |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
                :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
                |||| |||||||||:||||||||||||||||||||| ||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
                ||||| |||| ||| |||||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
                |||||| |||:|||||||||||| |||||||||||||| |||   |  |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
                |||| | ||  | ||       |||||||||||||:|||||||||||| ||:||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
                ||||   ||| | |      ||||| ||||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
                ||||:|||||| |||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
                ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
                ||||||| | |||| |||||||||||||| ||||||||||||||||||||||| ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
                |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
                |||||||||| ||| |||||||||||| |||||||        | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
                |||| |||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
                |||||||| |||||||||| | |||||||||||||  |||||||||||||||||||||
```

```
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
             ||||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
             |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
             |||||| ||||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 158
CS248653
LOCUS       CS248653               1511 bp    DNA     linear   PAT 09-JAN-2006
DEFINITION  Sequence 21 from Patent EP1612270.
ACCESSION   CS248653
VERSION     CS248653.1  GI:84662046
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Parahnos-Baccala,G.
  TITLE     Retroviral nucleic material and nucleotide fragments, in particular
            associated with multiple sclerosis and/or rheumatoid arthritis, for
            diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: EP 1612270-A 21 04-JAN-2006;
            BIOMERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1511
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:|||| |||| |||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC---------------------------------- 727
              | ||||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| |||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||| || |||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||||| |||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |   |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      |||||||||||||:||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:|||||||  |||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||| |||||||||||||||||||||| ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||  |:|| ||||||||||||||||||||||
Db        1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||||| |||||||      | ||||  |||||||
Db        1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||||  |||||||||||||||| ||||||||||||||||||||||
Db        1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db        1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||  ||||||||||||||||||||||
Db        1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||||  ||| |  | ||||||||||||| ||||||||||||||||||
Db        1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||||| |
Db        1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| ||||||||||||| ||||
Db        1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

```
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 159
AR344397
LOCUS       AR344397               1511 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 117 from patent US 6582703.
ACCESSION   AR344397
VERSION     AR344397.1  GI:33740338
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1511)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 117 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1511
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:|||||||:||||  ||||  |||||||  ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               |||||||||||||||||||||||||| ||||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||||| |||||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||||| |||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| ||||||||||:|||||||||||||||||||||| ||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |  |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||       |||||||||||||:|||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |      ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| ||||||||||:|||| ||: |||:|||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| ||||||||||||||| ||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||  |:|| ||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||||| ||||||||        | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||||| |||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | |||||||||||||| ||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 160
AX001039
LOCUS       AX001039               1511 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 21 from Patent WO9902666.
ACCESSION   AX001039
VERSION     AX001039.1  GI:7241274
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1511)
```

```
  AUTHORS   Ott,C. and Bedin,F.
  TITLE     RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR
            ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR
            DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
  JOURNAL   Patent: WO 9902666-A 21 21-JAN-1999;
            BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES             Location/Qualifiers
     source          1. .1511
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;

Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  ||||||||||||:||||||||:|||| |||| ||||||| ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | |||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               ||||||||||||||||||||||||||||| ||||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
              | |||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| |||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
              |||||     | |||||||||||||||| ||::|||||||||||||||| |||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||| ||||||||| |||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||||| |||| |||| |||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| |||||||||:|||||||||||||||||||| ||||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| ||||||||||||||| |||   |   |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      |||||||||||||:|||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| |||||||||:|||| ||: |||:|||||||||||||||||||||| |
Db        829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||||| |||||| |||||||| ||||||||||||||||||
Db        888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| |||||||||||||||  ||||||||||||||||||||||  ||||
Db        947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| ||||||||| |:|| ||||||||||||||||||||||||
Db       1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||| ||| |||||||||||| |||||||       | ||||  |||||||
Db       1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||  |||||||||||||||| ||||||||||||||||||||||||
Db       1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db       1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||| | |||||||||||||  |||||||||||||||||||||||
Db       1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||  ||| |  | |||||||||||||| ||||||||||||||||||||
Db       1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db       1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db       1487 GGTATGTTTGACCATTGAGAGCCA 1510


RESULT 161
AX027451
LOCUS       AX027451               1511 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 1 from Patent FR2788784.
ACCESSION   AX027451
VERSION     AX027451.1  GI:10188417
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 1 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1511
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             14.9%;  Score 1117.7;  DB 1;  Length 1511;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches 1327;  Conservative   18;  Mismatches  118;  Indels   61;  Gaps   12;
```

```
Qy       639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
Db         1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy       699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
Db        61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy       728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
Db       121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy       786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
Db       181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy       844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
Db       241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy       904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
Db       301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy       959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
Db       361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy      1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
Db       421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy      1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
Db       480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy      1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
Db       540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy      1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
Db       600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy      1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
Db       659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy      1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
Db       717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy      1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
Db       773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828

Qy      1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
Db       829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy      1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
Db       888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy      1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
Db       947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy      1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
Db      1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy      1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
Db      1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy      1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
Db      1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy      1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
Db      1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy      1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
Db      1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy      1912 AAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
Db      1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy      1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
Db      1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy      2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
Db      1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy      2092 GGTATGCTTGACCATTGAGGGCCA 2115
Db      1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

RESULT 162
AF123881
LOCUS AF123881 1511 bp mRNA linear VRL 11-AUG-1999
DEFINITION Multiple sclerosis associated retrovirus element gag polyprotein (gag) mRNA, partial cds.
ACCESSION AF123881
VERSION AF123881.1 GI:5726237
KEYWORDS .
SOURCE Multiple sclerosis associated retrovirus element
ORGANISM Multiple sclerosis associated retrovirus element
Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE 1 (bases 1 to 1511)
AUTHORS Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F., Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E., Mallet,F., Mandrand,B. and Perron,H.
TITLE Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles
JOURNAL Virology 260 (1), 1-9 (1999)
PUBMED 10405350
REFERENCE 2 (bases 1 to 1511)
AUTHORS Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F., Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E., Mallet,F., Mandrand,B. and Perron,H.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1999) UMR103, CNRS/Biomerieux, 46 Allee D'Italie, Lyon Cedex 07 69364, France
FEATURES Location/Qualifiers
source 1. .1511
/organism="Multiple sclerosis associated retrovirus element"
/mol_type="mRNA"
/specific_host="Homo sapiens"
/db_xref="taxon:89382"
/clone="CL2"
/note="multiple sclerosis patient cell-free plasma"
gene <1. .>1511
/gene="gag"
CDS <1. .1060
/gene="gag"

/note="encodes partial matrix and capsid proteins; gag polyprotein (CA) has 3 potential glycosylation sites"
/codon_start=2
/product="gag polyprotein"
/protein_id="AAD48375.1"
/db_xref="GI:5726238"
/translation="LERILENWDQCDTQTLRKKRFIFFCSTAWPQYPLQGRETWLPEG
SINYNIILQLDLFCRKEGKWSEVPYVQTFFSLRDNSQLCKKCGLCPTGSPQSPPPYPS
VPSPTPSSTNKDPPLTQTVQKEIDKGVNNEPKSANIPRLCPLQAVRGGEFGPARVPVP
FSLSDLKQIKIDLGKFSDNPDGYIDVLQGLGQSFDLTWRDIMLLLNQTLTPNERSAAV
TAAREFGDLWYLSQANNRMTTEERTTPTGQQAVPSVDPHWDTESEHGDWCHKHLLTCV
LEGLRKTRKKPMNYSMMSTITQGKEENLTAFLDRLREALRKHTSLSPDSIEGQLILKD
KFITQSAADIRKNFKSLP"

Query Match 14.9%; Score 1117.7; DB 1; Length 1511;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 1327; Conservative 18; Mismatches 118; Indels 61; Gaps 12;

```
Qy        639 CCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG 698
              ||||  |||||||||||:||||||||:||||  ||||  |||||||  ||||||||||||||
Db          1 CCTAGAACGTATTCTGGAGAATTGGGACCAATGTGACACTCAGACGCTAAGAAAGAAACG 60

Qy        699 ACTTATATTCTTCTGCAGTGCCGCCTGGC------------------------------ 727
              | ||||||||||||||||| |||||||||
Db         61 ATTTATATTCTTCTGCAGTACCGCCTGGCCACAATATCCTCTTCAAGGGAGAGAAACCTG 120

Qy        728 -ACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACYTCTTTTGTAGAA- 785
               ||||||||||||||||||||||||| ||||||||||||||:|||| |||||||
Db        121 GCTTCCTGAGGGAAGTATAAATTATAACATCATCTTACAGCTAGACCTCTTCTGTAGAAA 180

Qy        786 --AAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACAACTC 843
                | |||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        181 GGAGGGCAAATGGAGTGAAGTGCCATATGTGCAAACTTTCTTTTCATTAAGAGACAACTC 240

Qy        844 ACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTA 903
              |||||||||||||||||| |||||||||||||||||| |||||||| ||||||||
Db        241 ACAATTATGTAAAAAGTGTGGTTTATGCCCTACAGGAAGCCCTCAGAGTCCACCTCCCTA 300

Qy        904 TCCCAG-----CATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAAT 958
               |||||      | |||||||||||||||| ||::|||||||||||||||||| ||||||||
Db        301 CCCCAGCGTCCCCTCCCCGACTCCTTCCTCAACTAATAAGGACCCCCCTTTAACCCAAAC 360

Qy        959 GGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCA 1018
              |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db        361 GGTCCAAAAGGAGATAGACAAAGGGGTAAACAATGAACCAAAGAGTGCCAATATTCCCCG 420

Qy       1019 ATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGC 1078
              |||||| ||||| |||||||||| ||||| |||||||||||||||||||||| ||| |
Db        421 ATTATGCCCCCT-CCAAGCAGTGAGAGGAGGAGAATTCGGCCCAGCCAGAGTGCCTGTAC 479

Qy       1079 YTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYC 1138
              :||||::|||| |||||||||||||| |||| |||| ||||||||||||||||||:|
Db        480 CTTTTTCTCTCTCAGACTTAAAGCAAATTAAAATAGACCTAGGTAAATTCTCAGATAACC 539

Qy       1139 CTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGAG 1198
              |||| ||||||||||:|||||||||||||||||||||| ||||||||||||||||||||
Db        540 CTGACGGCTATATTGATGTTTTACAAGGGTTAGGACAATCCTTTGATCTGACATGGAGAG 599

Qy       1199 ATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAACT 1258
              ||||| |||| ||| |||||||||||||||||||||||||||||||||| |  |||||
Db        600 ATATA-ATGTTACTACTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCGCTGTAACT 658

Qy       1259 GCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAA 1318
              |||||| |||:|||||||||||| |||||||||||||||| |||   |  |||||||||
Db        659 GCAGCCCGAGAGTTTGGCGATCTTTGGTATCTCAGTCAGGCCAA--CAATAGGATGACAA 716

Qy       1319 CAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTASACCCTCAT 1378
              |||| | ||  | ||      ||||||||||||:|||||||||||| ||:|||||||||
Db        717 CAGAGGAAAGAACAA----CTCCCACAGGCCAGCAGGCAGTTCCCAGTGTAGACCCTCAT 772

Qy       1379 TGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGT 1438
              ||||   ||| | |     ||||| |||||||||||  || ||||||||||||||| ||
Db        773 TGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCCACAAACATTTGCTAACTTGCGT 828
```

```
Qy       1439 GCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCA 1498
              ||||:||||||| |||||||||||:|||| ||: |||:|||||||||||||||||||| |
Db         829 GCTAGAAGGACTGAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGATGTCCACTA 887

Qy       1499 TAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTG 1558
              ||||||| ||||| |||||||||||| |||||| |||||||| ||||||||||||||||||
Db         888 TAACACA-GGGAAAGGAAGAAAATCTTACTGCTTTTCTGGACAGACTAAGGGAGGCATTG 946

Qy       1559 AGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAG 1618
              ||||||| | |||| ||||||||||||||||  |||||||||||||||||||||||  ||||
Db         947 AGGAAGCATACCTCCCTGTCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAG 1006

Qy       1619 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGG 1678
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        1007 TTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCTTAGGCCCGG 1066

Qy       1679 AGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGG 1738
              |||| |||||||||||||||| |||||||| |:|| ||||||||||||||||||||||
Db        1067 AGCAGAACTTAGAAACCCTATTTAACTTGGCATCCTCAGTTTTTTATAATAGAGATCAGG 1126

Qy       1739 AGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAA-------GGCCACCGCTTTAGT 1791
              |||||||||||| ||| |||||||||||||| ||||||||       | ||||  |||||||
Db        1127 AGGAGCAGGCGAAACGGGACAAACGGGATAAAAAAAAAAGGGGGGGTCCACTACTTTAGT 1186

Qy       1792 CATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG 1851
              |||| |||||||||||  |||||||||||||||| |||||||||||||||||||||||||
Db        1187 CATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGCAAAAGGGAAAAGCTGGGCAAATCA 1246

Qy       1852 AATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCC 1911
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db        1247 AATGCCTAATAGGGCTGGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTATCC 1306

Qy       1912 AAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGC 1971
              |||||||| |||||||||||| | |||||||||||||  |||||||||||||||||||||
Db        1307 AAGTAGAAATAAGCCGCCCCCTTGTCCATGCCCCTTACGTCAAGGGAATCACTGGAAGGC 1366

Qy       1972 CCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCA 2031
              |||||||||||||||   ||| |  | ||||||||||||||| ||||||||||||||||||
Db        1367 CCACTGCCCCAGGGGATGAAGATACTCTGAGTCAGAAGCCATTAACCAGATGATCCAGCA 1426

Qy       2032 GCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGCCCTG 2091
              |||||||||||||||| ||||| ||||||| |||||||||||||||||||||||||| |
Db        1427 GCAGGACTGAGGGTGCCCGGGGCGAGCGCCAGCCCATGCCATCACCCTCACAGAGCCCCG 1486

Qy       2092 GGTATGCTTGACCATTGAGGGCCA 2115
              |||||| |||||||||||| ||||
Db        1487 GGTATGTTTGACCATTGAGAGCCA 1510
```

RESULT 163
BD196240
LOCUS BD196240 1136 bp DNA linear PAT 17-JUL-2003
DEFINITION Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.
ACCESSION BD196240
VERSION BD196240.1 GI:33006010
KEYWORDS JP 2002512530-A/6.
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 1136)
AUTHORS Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
TITLE Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders
JOURNAL Patent: JP 2002512530-A 6 23-APR-2002;
BIO MERIEUX
COMMENT OS Unidentified
PN JP 2002512530-A/6
PD 23-APR-2002
PF 06-JUL-1998 JP 1999508244
PR 07-JUL-1997 FR 97/08815
PI FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD

```
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .1136
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1136
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              ||||||||||||||:||||||||||||||||||||:||||||||||||:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:||||||||||||||||||:||||||:||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db         781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
              |||:||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db         841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:|||||||:||||||||||||||||||||||||||||||||||||||||||||
Db         961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 164
BD267482
LOCUS       BD267482                1136 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267482
VERSION     BD267482.1  GI:33077250
KEYWORDS    JP 2002534980-A/23.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1136)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 23 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/23
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .1136
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1136
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60
```

```
Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             ||||||||||||||:|||||||||||||||||||:||||||||||||:||||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:|||||||||:||||||||||||||||:||||||:||||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116
```

```
RESULT 165
AX000961
LOCUS       AX000961               1136 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 6 from Patent WO9902696.
ACCESSION   AX000961
VERSION     AX000961.1  GI:7241203
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1136)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 6 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .1136
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             ||||||||||||||:|||||||||||||||||||:||||||||||||:||||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:|||||||||:||||||||||||||||:||||||:||||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720
```

```
Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:|||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 166
AX027475
LOCUS       AX027475               1136 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 25 from Patent FR2788784.
ACCESSION   AX027475
VERSION     AX027475.1  GI:10188439
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 25 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1136
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
             |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
             ||||||||||||||:||||||||||||||||||||:||||||||||:||||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
             ||||||||:|||||||||:||||||||||||||||||:||||||:||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
             ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
             |||:||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
             |||||||:|||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 167
AF072508
LOCUS       AF072508               1136 bp    mRNA    linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens endogenous retrovirus W envelope protein mRNA, partial
            cds.
ACCESSION   AF072508
VERSION     AF072508.1  GI:4262295
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 1136)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 1136)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1136
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl.C4C5"
                     /tissue_type="placenta"
                     /clone_lib="Clontech 5'-stretch plus library, Cat number
                     HL5014a"
     repeat_region   1. .1136
                     /rpt_family="HERV-W"
     CDS             <1. .731
                     /codon_start=3
                     /product="envelope protein"
                     /protein_id="AAD14548.1"
                     /db_xref="GI:4262296"
                     /translation="MAIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALGTGIGGI
                     TTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDSLTAERGGTCL
                     FLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILPFLGPLA
                     AIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASPRSDVND
                     IKGTPPEEISAAQPLLRPNSAGSS"
     misc_feature    778. .788
                     /note="polypurine tract"
     3'UTR           790. .1116
                     /note="U3-R junction undetermined"
     polyA_signal    1098. .1103
     polyA_site      1116

  Query Match             14.8%;  Score 1106.8;  DB 1;  Length 1136;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1093;  Conservative   23;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||:||||||||||||||||:||||||||||:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTCGCTAACCGCTGAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:||||||||||||||||:||||||:||||||||||||||||
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAACGTAGAGCAGAAGAGCTTCGAAACACTG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        421 GACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        481 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 7066
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        541 TCTCTTCCAGAATCGAAGCTGTAAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTA 600

Qy       7067 AGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCA 7126
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        601 AGATCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCA 660

Qy       7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 7186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAG 720

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 780

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||:||||||
Db        781 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCC 840

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
              |||:||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        841 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCT 900

Qy       7367 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 7426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGC 960

Qy       7427 CAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 7486
              |||||||:|||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAATCATCTATTGCATGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGT 1020

Qy       7487 YTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 7546
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTT 1080

Qy       7547 TCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              ||||||||||||||||||||||||||||||:||||:
Db       1081 TCATGCTATTTCACTCTATTAAATCTTGCAGCTGCG 1116


RESULT 168
CQ745026
LOCUS       CQ745026                1344 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30960 from Patent WO02068579.
ACCESSION   CQ745026
VERSION     CQ745026.1  GI:42362214
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 30960 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
  Query Match           14.1%;  Score 1054.1;  DB 1;  Length 1344;
  Best Local Similarity  85.5%;  Pred. No. 0;
  Matches 1160;  Conservative   82;  Mismatches   94;  Indels   21;  Gaps    9;

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
             | |||||||:||||:|||||| ||||:|||||:|||||:|| ||||||||||||||||||
Db             6 TTCCCAGGTATGGCGAAATAG-CAGGCCATTATATACACTAGTTAAGGAAACTCAGAAAG 64

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
             |||||  |||||||:||| |||||:|:||||||:::|||||||||||||||||||  ||
Db            65 TCAATAACCATTTAGTAAAATGGACACCTGAAGCAAAAGTGGCTTTCCAGGCCCTAAAGA 124

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
             | | |||  ||||||||:|||||||||||:||||||||:||||||||||:||:| |
Db           125 AGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTGTTTAAA 182

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             ::|||||| |||||||||:||:|||||||:|||| ||||||||||:|||:|||||:|||||
Db           183 TGTCACAG-AAAAAACAGGAATAGCTCTAGGAGACCTTACACAGGTCCAAGGGATGAGCT 241

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:|||||
Db           242 TGCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATT 301

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||: |||||||| |||||||||||| :|||||||| || ||||||||||||||||||
Db           302 GTTTACTGGTAGTGGCGGCAGTAGCAGTGTTAGTATCTAAAACAGTTAAAATAATACAGG 361

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |: |||||||||||||||||| |||||||||:||||::|||||||||||||||||||||||
Db           362 GAGGAGATCTTACTGTGTGGATATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 421

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAG 3748
             |||||||||||||||||||:|||||||||||||:||||||||||||||||||||:||||||
Db           422 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAG 481

Qy       3749 TGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAG 3808
             ||||||:|||| |||||| |||||||||||||||||:| ||||||||||||||||||||||
Db           482 TGCTGCGACTGCGCACTTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAG 541

Qy       3809 AAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACC 3868
             |||| |||:||:|||||||   |||:| ||| | |||||||  ||| ||||||||||
Db           542 AAAAAATAGAACATAACTGT---CAAGTGATTGCCCAAACCTACACCATTCGAGGGGACC 598

Qy       3869 TTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTT 3928
             ||:||||:|||||:|||||||||||: ||| |||||||| ||||||||||||||||||||
Db           599 TTCTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTCTACTGATGGAAGTTCCTTT 657

Qy       3929 GTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAA 3988
             |||||||||||||||| |||||: |||||||||||||||| ||||||||||:|| ||
Db           658 GTAGAAAAAGGACTTCAAAAAGCAGGGTATGCAGTGGTCAGTAATAATGGAATACTTAAA 717

Qy       3989 AGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGG 4048
             |||||||||||||||||| ||||||||||:||||:|||||||||||||||:|||:|  |
Db           718 AGTAATCCCCTCACTCCAGAAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCAG 777

Qy       4049 GCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYAC 4108
             |||||||||||||||| || :|||||||  :|||||||||||||||||:|:||||||:||
Db           778 GCACTAGAATTAGGAAAAAAAAAAAGGATAAATATATATACAGACTCTAAGTATGCTTAC 837

Qy       4109 CTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGA 4168
             |||||| |||||||| |||::|||||||:|:|||||| |||||||| |||||:|||: |
Db           838 CTAGTCCTCCATGCCAATGCAGCAATATGGAGAGAAAGAGAATTCCTACCTTCTGAGGAA 897

Qy       4169 ACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARA 4228
             ||||||||||:|| |||||||| |||||||:| |||||:||||:||||| |||||||:|
Db           898 ACACCTATCAAACGTCAGGAAGCGATTAGGAGA-TATTATTGGCTGTACAAAAACCTAAA 956

Qy       4229 GAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRG 4288
             ||         |||||||| |:|| ||||||| |||  ||| :||||||||||:|||:
Db           957 GA--------CCTTACACTACCGGAGTCATCAGAAAAAAAAAAAAAGGGAAATAGAAGGA 1008

Qy       4289 AAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATG 4348
             ||: ||||||| :||||| ||||:|||||||  |||||||||||||||||| ||||||
Db          1009 AACCGCCAAGCGGATATTAAAGCCAAAAGAGCCACAAGGCAGGACCCTCCATTAAAAATG 1068

Qy       4349 CTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCA 4408
             |||||  ||   ||| || ||| || ||||||| |||||||||||||||||||||||||||
Db          1069 CTTATAGAAGGACCCCTAATATGGGATAATCCCCTCCGGGAAACCAAGCCCCAGTACTCA 1128

Qy       4409 GCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCCTCGGGACGGTTA 4465
             || | ||| |||||||||||||||||||||     ||||| |||||||| ||| || ||
Db          1129 AAAGAAAAAATAGAATGGGGAACCTCACGAGGACATAGTTTCCTCCCCTCAGGATGGCTA 1188

Qy       4466 GCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACC 4525
             |||||  ||||||| |||||||||| ||| ||| ||| ||||||||||||||||||||||
Db          1189 GCCACCAAAGAAGGAAAAATACTTTCGCCAGCAGCTAACCAATGGAAATTACTTAAAACC 1248

Qy       4526 CTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATTT 4585
             ||||| |||||||||||||||| || ||||||| ||||||:|||||||||||||||||||
Db          1249 CTTCACCAAACCTTTCACTTAGG-ATTGATAGCAACCATCAGATGGCCAAATCATTATTT 1307

Qy       4586 ACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAK 4622
             ||||||||||||||||||||||||||||||||||:|||:
Db          1308 ACTGGACCAGGCCTTTTCAAAACTATCAAGCAGATAG 1344


RESULT 169
BD221828
LOCUS       BD221828                1080 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221828
VERSION     BD221828.1  GI:33031598
KEYWORDS    JP 2002518051-A/23.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 23 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002518051-A/23
            PD   25-JUN-2002
            PF   23-JUN-1999 JP 2000556036
            PR   23-JUN-1998 FR     98/07920
            PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
            C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
            A61P25/00,
            PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
            sequence and deduced protein sequence family with  CC
            human endogenous
            CC   retroviral motifs, and their uses
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(1080).
FEATURES              Location/Qualifiers
     source           1. .1080
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"

  Query Match           12.8%;  Score 960.2;  DB 1;  Length 1080;
  Best Local Similarity  96.4%;  Pred. No. 0;
  Matches 1052;  Conservative   16;  Mismatches   11;  Indels   12;  Gaps    9;

Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
              ||||||||||||||||||||||||||||||||||||||||::||||||||||||||||||
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
              ||||||||:||||::||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
              ||||||:|||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||     |
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
              |||||||||||| |||||||| ||||||||||||||||||||:|||||||||||||||:
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
              |||||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
              ||||||||||:|||||||||||||||:|||| ||:||||:|||||||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
              ||||||||||
Db       1070 TGAATGCCTAA 1080


RESULT 170
AR699777

LOCUS       AR699777               1080 bp    DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 28 from patent US 6919438.
ACCESSION   AR699777
VERSION     AR699777.1  GI:75205568
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 28 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             12.8%;  Score 960.2;  DB 1;  Length 1080;
  Best Local Similarity   96.4%;  Pred. No. 0;
  Matches 1052;  Conservative   16;  Mismatches   11;  Indels   12;  Gaps    9;

Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
              ||||||||||||||||||||||||||||||||||||||||::||||||||||||||||||
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
              ||||||||:||||::||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
              ||||||:|||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||     |
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
              |||||||||||| |||||||| ||||||||||||||||||||:|||||||||||||||:
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
              |||||||||||   ||| | |     ||||| ||||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
              ||||||||||:|||||||||||||||:|||| ||:||||:|||||||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710
```

```
Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
             ||||||||||||
Db       1070 TGAATGCCTAA 1080
```

```
RESULT 171
AX008005
LOCUS       AX008005               1080 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 28 from Patent WO9967395.
ACCESSION   AX008005
VERSION     AX008005.1  GI:9995698
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 28 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1. .63
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07393.1"
                     /db_xref="GI:9995699"
                     /translation="TSFVEKANGVKCHKYKLSFH"

  Query Match            12.8%;  Score 960.2;  DB 1;  Length 1080;
  Best Local Similarity   96.4%;  Pred. No. 0;
  Matches 1052;  Conservative   16;  Mismatches   11;  Indels   12;  Gaps    9;
```

```
Qy        771 ACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 830
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAT 60

Qy        831 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGA 120

Qy        891 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCA 950
              ||||||||||||||||||||||||||||||||||||||::|||||||||||||||||||
Db        121 GTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCA 180

Qy        951 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAT 240

Qy       1011 ATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 1070
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGT 299

Qy       1071 GCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 1130
              ||||||||:|||||::||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTC 359

Qy       1131 AGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 1190
              ||||||:|||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        360 AGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGACAATTCTTTGATCTGAC 419

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATGGAGAGATATA-ATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 478

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||  |
Db        479 CCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAG 536

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
              |||||||||||| |||||||| ||||||||||||||||||||:|||||||||||||||:
Db        537 GATGACAACAG-AGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGTCTAG 595

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
              |||||||||||| ||| | |     ||||| ||||||||||||||||||||||||||||||
Db        596 ACCCTCATTGGG---ACACAGAATCAGAACAT-GGAGATTGGTGCTGCAGACATTTGCTA 651

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
              ||||||||||||:|||||||||||||||||||:|||| ||:||||:|||||||||||||||
Db        652 ACTTGTGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTACTCAATGAT 710

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        711 GTCCACCATAACACA-GGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 769

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 829

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        830 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCG 889

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        890 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACCTCGGTTTTTTATAATA 949

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1009

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1069

Qy       1850 TGAATGCCTAA 1860
              ||||||||||||
Db       1070 TGAATGCCTAA 1080
```

```
RESULT 172
A60173
LOCUS       A60173                 1167 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 61 from Patent WO9706260.
ACCESSION   A60173
```

```
VERSION     A60173.1  GI:3715184
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 61 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db        119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db        177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db        356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db        416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db        476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db        536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db        595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db        1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db        1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db        1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 173
A79523
LOCUS       A79523                  1167 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 61 from Patent EP0789077.
ACCESSION   A79523
VERSION     A79523.1  GI:6092526
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1167)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 61 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db        119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db        177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236
```

```
Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db        356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db        416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db        476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db        536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db        595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167

RESULT 174
AR094748
LOCUS       AR094748               1167 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 61 from patent US 6001987.
ACCESSION   AR094748
VERSION     AR094748.1  GI:10021961
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1167)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 61 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db        119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db        177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db        356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db        416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db        476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db        536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db        595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834
```

```
Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 175
AR343195
LOCUS       AR343195               1167 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 61 from patent US 6579526.
ACCESSION   AR343195
VERSION     AR343195.1  GI:33738697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1167)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 61 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db        119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db        177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db        237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db        296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db        356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db        416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db        476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db        536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db        595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db        655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db        715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db        775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db        835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db       1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db       1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 176
AR344348
LOCUS       AR344348               1167 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 57 from patent US 6582703.
ACCESSION   AR344348
VERSION     AR344348.1  GI:33740289
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1167)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
```

```
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 57 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             12.7%;  Score 952.8;  DB 1;  Length 1167;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches 1013;  Conservative   77;  Mismatches   73;  Indels   10;  Gaps    6;

Qy       3313 AAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCT 3372
Db          1 AAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCAGAAGCAGAAGCAGCT 60

Qy       3373 TTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGC 3432
Db         61 TTCCAGGCCCTAAAGAAATCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGC 118

Qy       3433 AAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAG 3492
Db         119 AAGACTTTTCTTTATATGTCACAG--AAAAACAGGAATAGCTCTAGGAGTCCTTACACAG 176

Qy       3493 RTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCA 3552
Db         177 GTCCAAGGGACAAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTANTGGCA 236

Qy       3553 AAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCA 3612
Db         237 AAGGGTTGGCCTCATTGTTTACAGGTAG-GGCAGCAGTAGCAGTCTTAGTTTCTGAAACA 295

Qy       3613 GTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA 3672
Db         296 GTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATA 355

Qy       3673 CTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTA 3732
Db         356 CTCACTGCTAAAGAGGACTTGTGGCTGTCAGACAACCATTTACTTAAATAGCAGGTTCTA 415

Qy       3733 TTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTT 3792
Db         416 TTACTTGAAGTGCCAGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTT 475

Qy       3793 CTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTAT 3852
Db         476 CTTCCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTAT 535

Qy       3853 GCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATAC 3912
Db         536 GCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATAC 594

Qy       3913 TGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGA 3972
Db         595 TGATGGAAGTTCCTTGGCAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGATCAGTGA 654

Qy       3973 TAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACT 4032
Db         655 TAATGGAATACTTGAAAGTAATCGCCTCACTCCAGGAACTAGTGCTCACCTGGCAGAACT 714

Qy       4033 AATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGA 4092
Db         715 AATAGCCCTCACTTGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATTCAGA 774

Qy       4093 CTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATT 4152
Db         775 CTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAGAGGGAATT 834

Qy       4153 CCTAACTTCYGAGRGAACACCTATCAMACATCA-GGAAGCCATTAGGARATTATTAYTGG 4211
Db         835 CCTAACTTCTGAGGGAACACCTATCAACCATCAGGGAAGCCATTAGGAGATTATTATTGG 894

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db         895 CTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCAGGGTCATCAGGAAGAAGAGG 954

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db         955 AAAGGGAAATAGAAGGCAATCGCCAAGCGGATATTGAAGCAAAAAAAGCCGCAAGGCAGG 1014

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db         1015 ACTCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAA 1074

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTC 4448
Db         1075 CCAAGCCCCAGTACTCAGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCC 1134

Qy       4449 TCCCCTCGGGACGGTTAGCCACTGAAGAAGGGA 4481
Db         1135 TCCCCTCCAGATGGCTAGCCACTGAGGAAGGAA 1167


RESULT 177
CQ082196
LOCUS       CQ082196                1117 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 17996 from Patent WO0157278.
ACCESSION   CQ082196
VERSION     CQ082196.1  GI:41052065
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 17996 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN HELA, SIGNAL = 1.3
                     NT HIT: AB026898.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P26808, EVALUE 4.00e-64"

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
Db            61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
Db         119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
Db         178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
```

```
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 178
CQ116881
LOCUS       CQ116881               1117 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25740 from Patent WO0157272.
ACCESSION   CQ116881
VERSION     CQ116881.1  GI:41086751
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 25740 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 6
                     NT HIT: AB026898.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P26808, EVALUE 4.00e-64"

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
Db        119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
Db        178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
Db        238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
Db        298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
Db        358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
Db        418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
Db        478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
Db        538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
Db        597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
Db        657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
Db        717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
```

```
                |||||:|:||||||:|||||||| ||||||||||::|||||||:|:|||||||| ||
Db        774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
                |||||||||:|||:|||||||||||:||||||||||||||||||:||||||:|||
Db        834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
                | |||||||||||||:||||||:|||||||||||||:|||||||| ||||||||:
Db        894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
                ||||||||||:|||:|||:|||||||| :| ||||||||:||||||||||||||||||
Db        954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
                |||||||||||||||||||| ||    ||| |||||| |||||||| || ||||||
Db       1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
                |||||||||||||||||  || ||||| ||||||||||||||||
Db       1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 179
CQ238747
LOCUS       CQ238747               1117 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25586 from Patent WO0157273.
ACCESSION   CQ238747
VERSION     CQ238747.1  GI:41222138
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 25586 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2
                     NT HIT: AB026898.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P26808, EVALUE 4.00e-64"

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
                |||||||||||||||||||||||||||:|||||||||:|:|||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
                ||||||||  ||  | | |||  |||||||||:||||||||   |||||:|||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
                |||||||||  :||:|||::|||||| |||||||||:||:||||||:|||||||||||||
Db         119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
                |:|||:|||||: ||| |||||||::||||::|||||:|||||||:||||||||||||
Db         178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
                |||||||||:|:|||||||||: |||||||| |||||||||||||::|||||||||||||
Db         238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
                ||||||| |||||||||  :|||||||||||||||||||||||||||||||:|||:|::||||
Db         298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
                |||||||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db         358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
                | | ||||||:| ||||||||||:|||||||||||||||||||||||||||||:| ||||
Db         418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
                |||||||| |||||||||||||||| ||:|||||||||||||:|||||  |||||||||
Db         478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
                |||||||  ||||||||||:||||:|||||: |||||||||||:|||| ||||||||||||
Db         538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
                |||||||||||||||| |||||||||| |||| |||||:|||||||||||| |||||||
Db         597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
                ||||||||||:|||||||||||||||||||||||||||||||||||||:||||:|||||||
Db         657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
                ||||||||:|||:|  ||||  |||||||  |||||:||||||:|||||||||||||||
Db         717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
                |||||:|:||||||:|||||||| ||||||||||::||||||||:|:|||||||| ||
Db         774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
                ||||||||||:|||:|||||||||||:|||||||||||||||||||:|||||||:|||
Db         834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
                | |||||||||||||:||||||||:|||||||||||||:|||||||||| ||||||||:
Db         894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
                ||||||||||:|||:|||:|||||||| :| ||||||||:||||||||||||||||||
Db         954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
                ||||||||||||||||||||  ||   ||| |||||| |||||||| || ||||||
Db        1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
                |||||||||||||||||  || ||||| ||||||||||||||||
Db        1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 180
CQ276425
LOCUS       CQ276425               1117 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24686 from Patent WO0157277.
ACCESSION   CQ276425
```

```
VERSION     CQ276425.1  GI:41249029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 24686 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
                     NT HIT: AB026898.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P26808, EVALUE 4.00e-64"

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  ||||||||:||||||||    ||||:||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118

Qy       3432 CAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACA 3491
             |||||||||  :||:|||::|||||| |||||||||:||:|||||||:|||||||||||||
Db         119 CAAGACTTTCCTTTATATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACA 177

Qy       3492 GRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGC 3551
             |:|||:|||||: ||| |||||||::||||::|||||||:||||||||:||||||||||||
Db         178 GGTCCAAGGGACAAGCCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGC 237

Qy       3552 AAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGC 3611
             |||||||||:|:|||||||||| |||||||| ||||||||||||::||||||||||||||
Db         238 AAAGGGTTGGCCTCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGC 297

Qy       3612 AGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAT 3671
             ||||||| ||||||||| :|||||||||||||||||||||||||||||||||:||||::||||
Db         298 AGTTAAAGTAATACAGGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCAT 357

Qy       3672 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCT 3731
             |||||||||||||||||||||||||||||||||||:||||||||||||:|||||||||
Db         358 ACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCT 417

Qy       3732 ATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATT 3791
             | | ||||||:| ||||||||||:|||||||||||||||||||||||||||:| ||||
Db         418 ACTGCTTGAAGGACCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATT 477

Qy       3792 TCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTA 3851
             |||||||| |||||||||||||||| ||:|||||||||||:|||||  |||||||||
Db         478 TCTTCCAGGCAATGAAGAAAAGATATAACATAACTGTCAACAAGTAATTGTTCAAACCTA 537

Qy       3852 TGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATA 3911
             ||||||| ||||||||||:||||:|||||: ||||||||||:|||| ||||||||||||
Db         538 CGCCACTCAAGGGGACCTTTTAGAGGTTCCCCTGACTGATCCCGACC-TCAACTTGTATA 596

Qy       3912 CTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTG 3971
             |||||||||||||||| |||||||||| |||| |||||:|||||||||||| |||||||
Db         597 CTGATGGAAGTTCCTTCGTAGAAAAAGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTG 656

Qy       3972 ATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAAC 4031
             ||||||||||:|||||||||||||||||||||||||||||||||:||||:|||||||
Db         657 ATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAC 716

Qy       4032 TAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAG 4091
             ||||||||:|||:|  |||| |||||||   |||||:||||||||:|||||||||||||||
Db         717 TAATAGCCCTCATTCAGGCAGTAGAATT---AGAAGGAAAAAGGGTAAATATATATACAG 773

Qy       4092 ACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAAT 4151
             |||||:|:|||||||:|||||||| ||||||||||||::||||||||:|:|||||||| ||
Db         774 ACTCTAAGTATGCTTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGGAT 833

Qy       4152 TCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGG 4211
             ||||||||||:|||:||||||||||||:|||||||||||||||||||||:|||||||:||
Db         834 TCCTAACTTCCGAGGGAACACCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGG 893

Qy       4212 CWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGR 4271
             | ||||||||||||||:|||||||||:|||||||||||||:||||||||| |||||||||:
Db         894 CGGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGG 953

Qy       4272 AAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGG 4331
             ||||||||||:|||:|||:|||||||| :| ||||||||:|||||||||||||||||||||
Db         954 AAAGGGAAATAGAAGGGAACTGCCAAGCGGACATTGAAGCAAAAAGAGCTGCAAGGCAGG 1013

Qy       4332 ACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAA 4391
             |||||||||||||||||||  ||  ||| ||||||  |||||||| || ||||||
Db         1014 ACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAA 1073

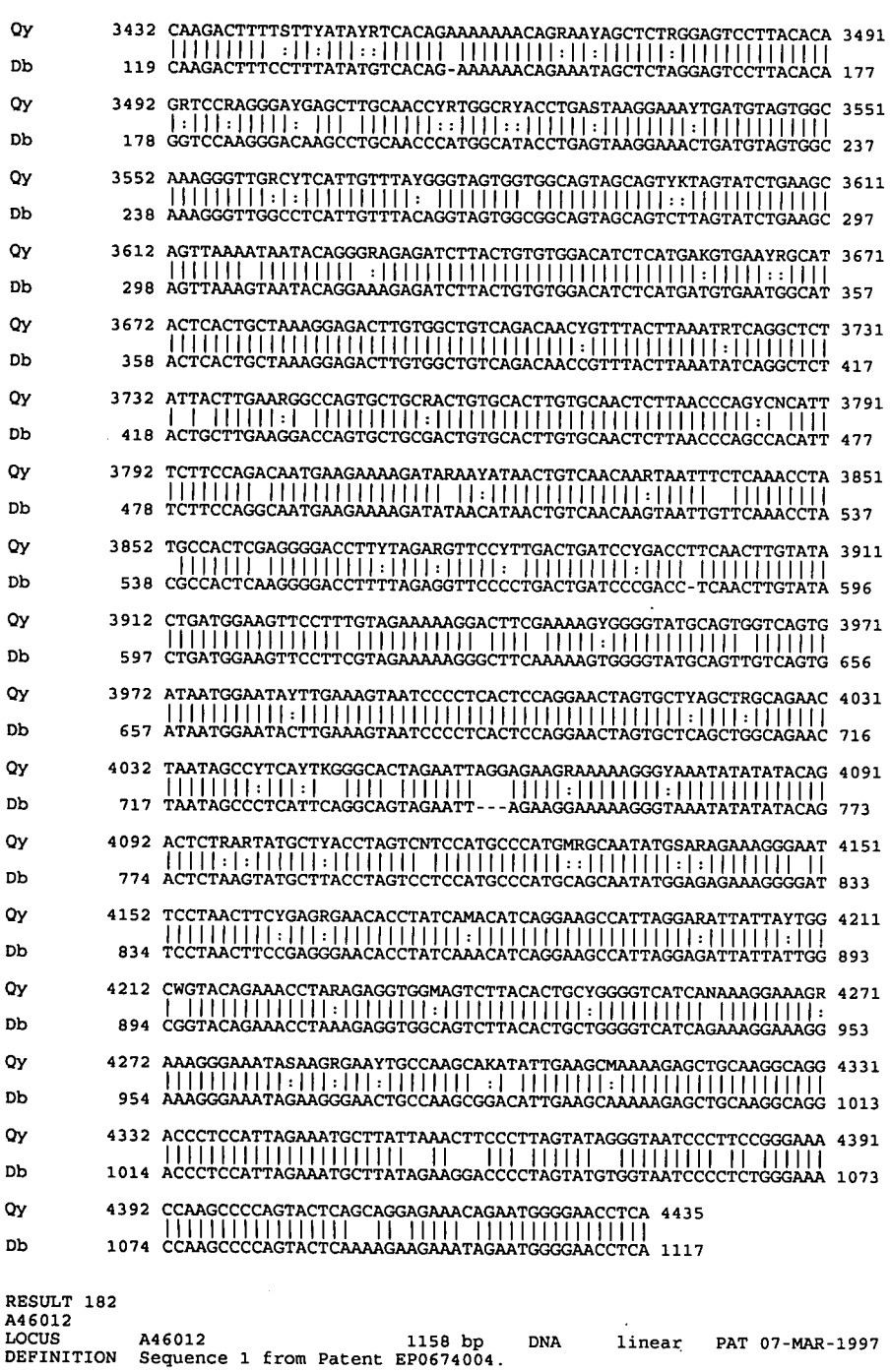
Qy       4392 CCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCA 4435
             |||||||||||||||||  || ||||| ||||||||||||||||
Db         1074 CCAAGCCCCAGTACTCAAAAGAAGAAATAGAATGGGGAACCTCA 1117


RESULT 181
CQ313425
LOCUS       CQ313425               1117 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24530 from Patent WO0186003.
ACCESSION   CQ313425
VERSION     CQ313425.1  GI:41274002
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 24530 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN LUNG, SIGNAL = 1.4
                     NT HIT: AB026898.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P26808, EVALUE 4.00e-64"

  Query Match             12.7%;  Score 948;  DB 1;  Length 1117;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  988;  Conservative   74;  Mismatches   52;  Indels   10;  Gaps    5;

Qy       3315 GGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTT 3374
             ||||||||||||||||||||||||||||||:|||||||||:|:||||||:::||| ||||||
Db          1 GGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTT 60

Qy       3375 CCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGG 3431
             ||||||||  ||  | | |||  ||||||||:||||||||    ||||:||||||||:|||
Db         61 CCAGGCCCTGAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGG 118
```



RESULT 182
A46012
LOCUS A46012 1158 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 1 from Patent EP0674004.

ACCESSION A46012
VERSION A46012.1 GI:2300284
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 1158)
AUTHORS Mandrand,B., Perron,H., Mallet,F., Bedin,F. and Beseme,F.
TITLE MSRV1 virus and pathogenic and/or infections agent MSRV2 linked to multiple sclerosis, their nucleic components and their applications
JOURNAL Patent: EP 0674004-A 1 27-SEP-1995;
BIO MERIEUX (FR)
COMMENT Other publication DE 674004T 960919
Other publication NO 953925 951204
Other publication FI 954699 951003
Other publication AU 1711495 950821
Other publication CA 2141907 950805
Other publication FR 2715939 950811
Other publication FR 2715938 950811
Other publication FR 2715937 950811
Other publication FR 2715936 950811.
FEATURES Location/Qualifiers
source 1. .1158
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 12.5%; Score 936.9; DB 1; Length 1158;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 990; Conservative 80; Mismatches 69; Indels 9; Gaps 6;

```
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||||| |||||:||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:|||| 
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||||::||| ::||||||:|||| 
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||:|:|||| |||||:||||| ||| |||||||||| 
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| ||||||||||||||| 
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: |||| 
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| |||||||||||||| 
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 183
A57048
LOCUS       A57048                  1158 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 1 from Patent WO9628552.
ACCESSION   A57048
VERSION     A57048.1  GI:3713031
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 1 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  |||||||| |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             ||||||||||||||||||||||||||||:|||||||||||:|||||||||  |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             ||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| |||||||  |:|||||||||||||:|||||||||||:|:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||||| |||||:||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:|||| 
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||||::||| ::||||||:|||| 
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||:|:|||| |||||:||||| ||| |||||||||| 
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||||| ||||||||||||||| 
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: |||| 
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| |||||||||||||| 
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149
```

```
Qy       3775 CTTAACCC 3782
              ||||| ||
Db       1150 CTTAAACC 1157


RESULT 184
A60113
LOCUS       A60113                  1158 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 1 from Patent WO9706260.
ACCESSION   A60113
VERSION     A60113.1  GI:3715129
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 1 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||    |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              |||| ||||| ||||| | |·|||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              ||||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              ||||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
              |||||||||||||||||:|||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
              |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
              |||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
              ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
              ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
              ||:::|||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db         617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||:|||||:|||||| |||||||||||||||||||||| |||||:||||||||:|
Db         676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||    :::|||||||||||||||||  ||  | | |||  ||||||||:|||||||||
Db         736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
              | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db         791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:|||||||||||| |||:||  |||||:|||||||||||||::||| ::||||||:||||
Db         850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:|||| |||||:|||| ||| ||||||||||
Db         910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||:: |||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db         970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:||||||::|||||||||||||||||||||||||||||||| ||||||||||: ||||
Db        1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db        1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
              ||||| ||
Db        1150 CTTAAACC 1157


RESULT 185
A79463
LOCUS       A79463                  1158 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 1 from Patent EP0789077.
ACCESSION   A79463
VERSION     A79463.1  GI:6092471
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 1 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||    |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
```

```
Db        78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy      2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814

Db       137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy      2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874

Db       197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy      2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934

Db       257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy      2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994

Db       317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy      2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054

Db       377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy      3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114

Db       437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy      3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174

Db       497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy      3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234

Db       557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy      3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294

Db       617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy      3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354

Db       676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy      3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414

Db       736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy      3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474

Db       791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy      3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534

Db       850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy      3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594

Db       910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy      3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654

Db       970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy      3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714

Db      1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy      3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774

Db      1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy      3775 CTTAACCC 3782

Db      1150 CTTAAACC 1157
```

RESULT 186
A80197
LOCUS A80197 1158 bp DNA linear PAT 20-OCT-1999
DEFINITION Sequence 1 from Patent WO9521256.
ACCESSION A80197
VERSION A80197.1 GI:6092929
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 1158)
AUTHORS Perron,H. and Mallet,F.
TITLE MSRV1 VIRUS AND MSRV2 PATHOGENIC AND/OR INFECTIOUS AGENT ASSOCIATED WITH MULTIPLE SCLEROSIS, NUCLEIC ACID COMPONENTS AND APPLICATIONS OF SAME
JOURNAL Patent: WO 9521256-A 1 10-AUG-1995;
BIO MERIEUX (FR); PERRON HERVE (FR)
FEATURES Location/Qualifiers
source 1. .1158
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 12.5%; Score 936.9; DB 1; Length 1158;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 990; Conservative 80; Mismatches 69; Indels 9; Gaps 6;

```
Qy      2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694

Db        19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy      2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754

Db        78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy      2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814

Db       137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy      2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874

Db       197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy      2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934

Db       257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy      2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994

Db       317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy      2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054

Db       377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy      3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114

Db       437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy      3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174

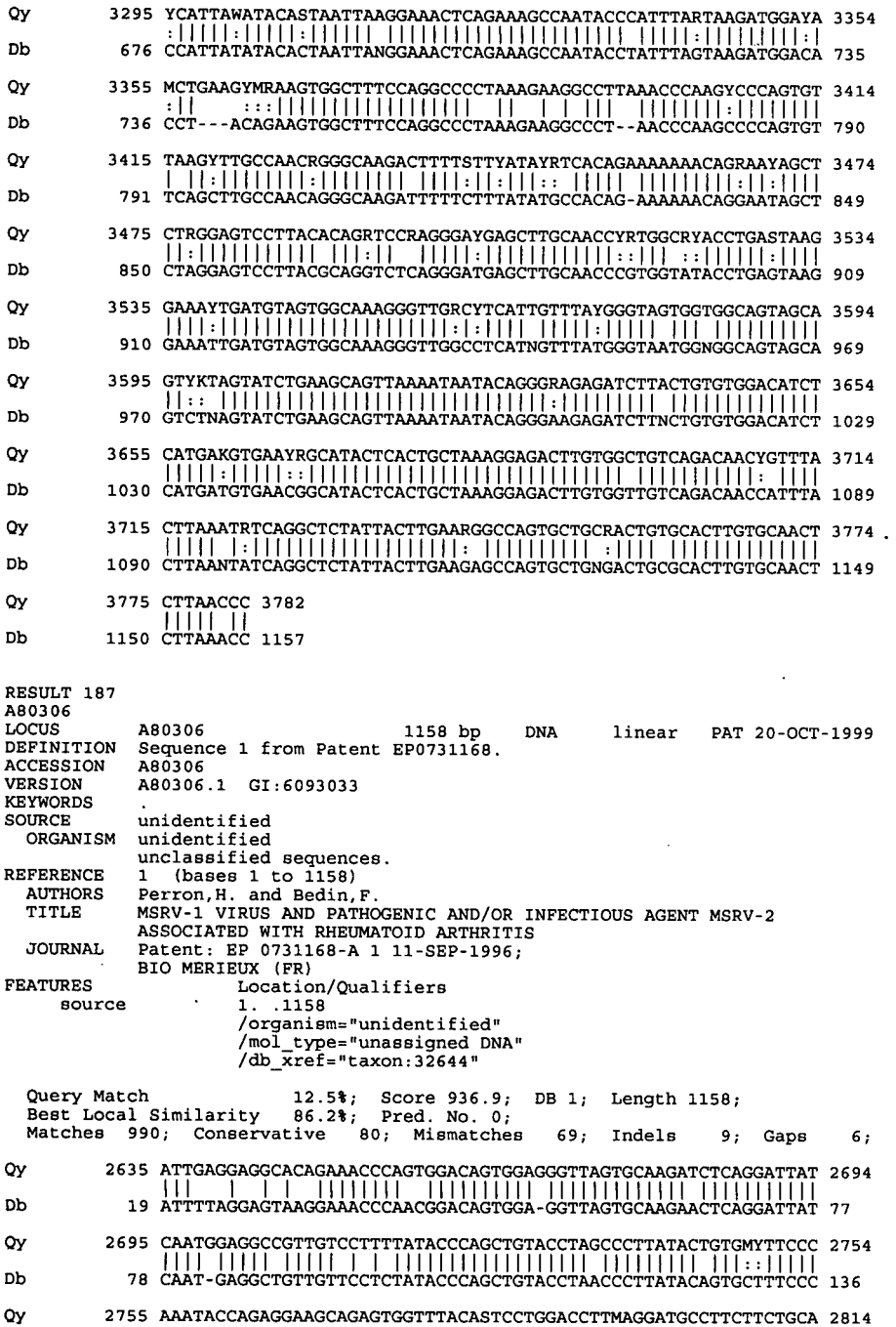
Db       497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy      3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234

Db       557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy      3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294

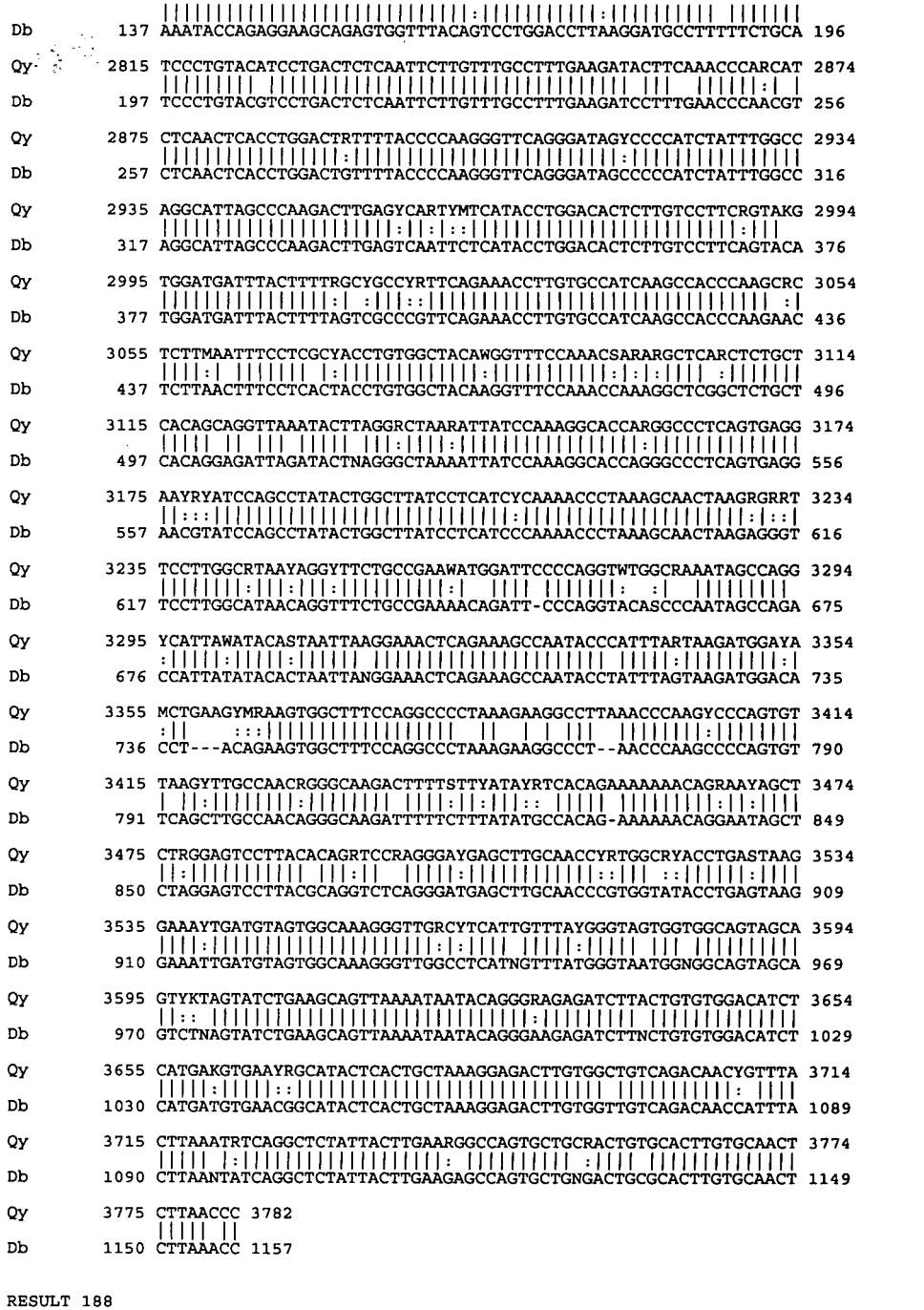
Db       617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675
```



```
RESULT 187
A80306
LOCUS       A80306                 1158 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 1 from Patent EP0731168.
ACCESSION   A80306
VERSION     A80306.1  GI:6093033
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 1 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;
```





```
RESULT 188
```

```
AR035349
LOCUS       AR035349               1158 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5871745.
ACCESSION   AR035349
VERSION     AR035349.1  GI:5952017
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Multiple sclerosis related virus
  JOURNAL   Patent: US 5871745-A 1 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db        1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db        1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db        1150 CTTAAACC 1157


RESULT 189
AR036190
LOCUS       AR036190               1158 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5871996.
ACCESSION   AR036190
VERSION     AR036190.1  GI:5952858
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Retrovirus agents MSRV1 and MSRV2 associated with multiple
            sclerosis
  JOURNAL   Patent: US 5871996-A 1 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

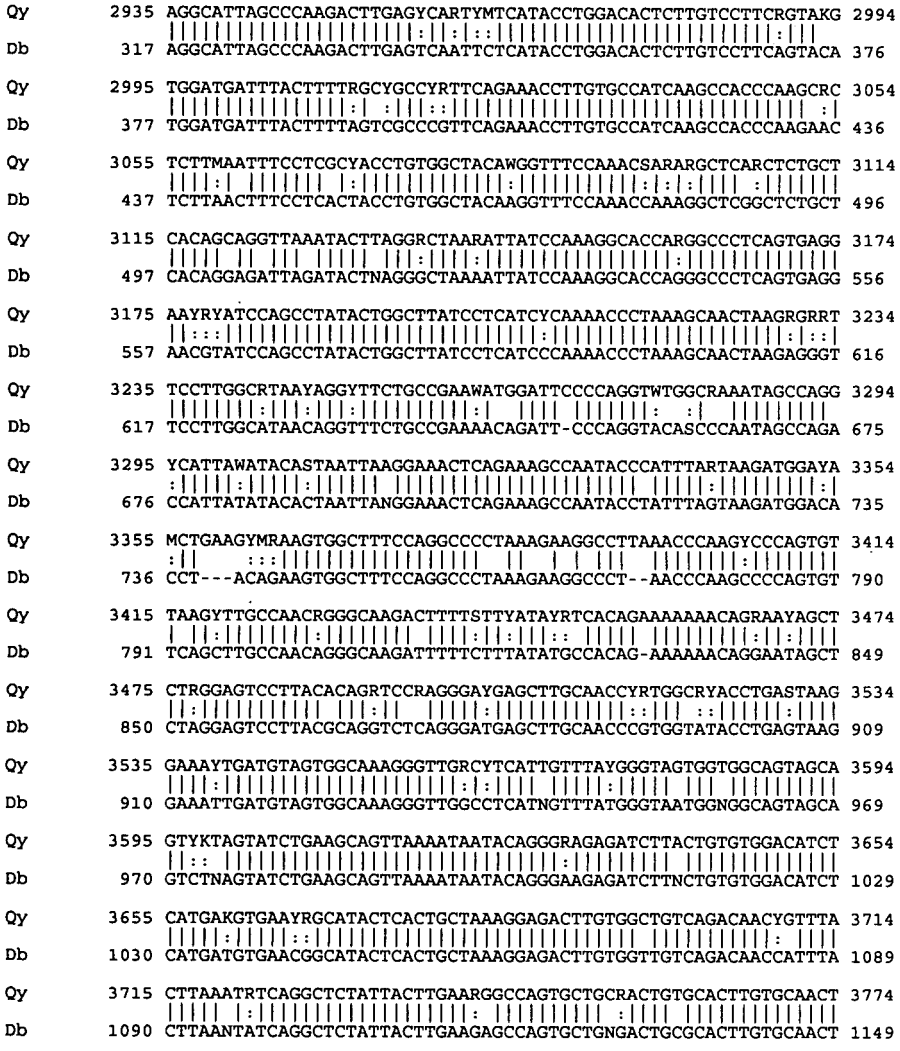
Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

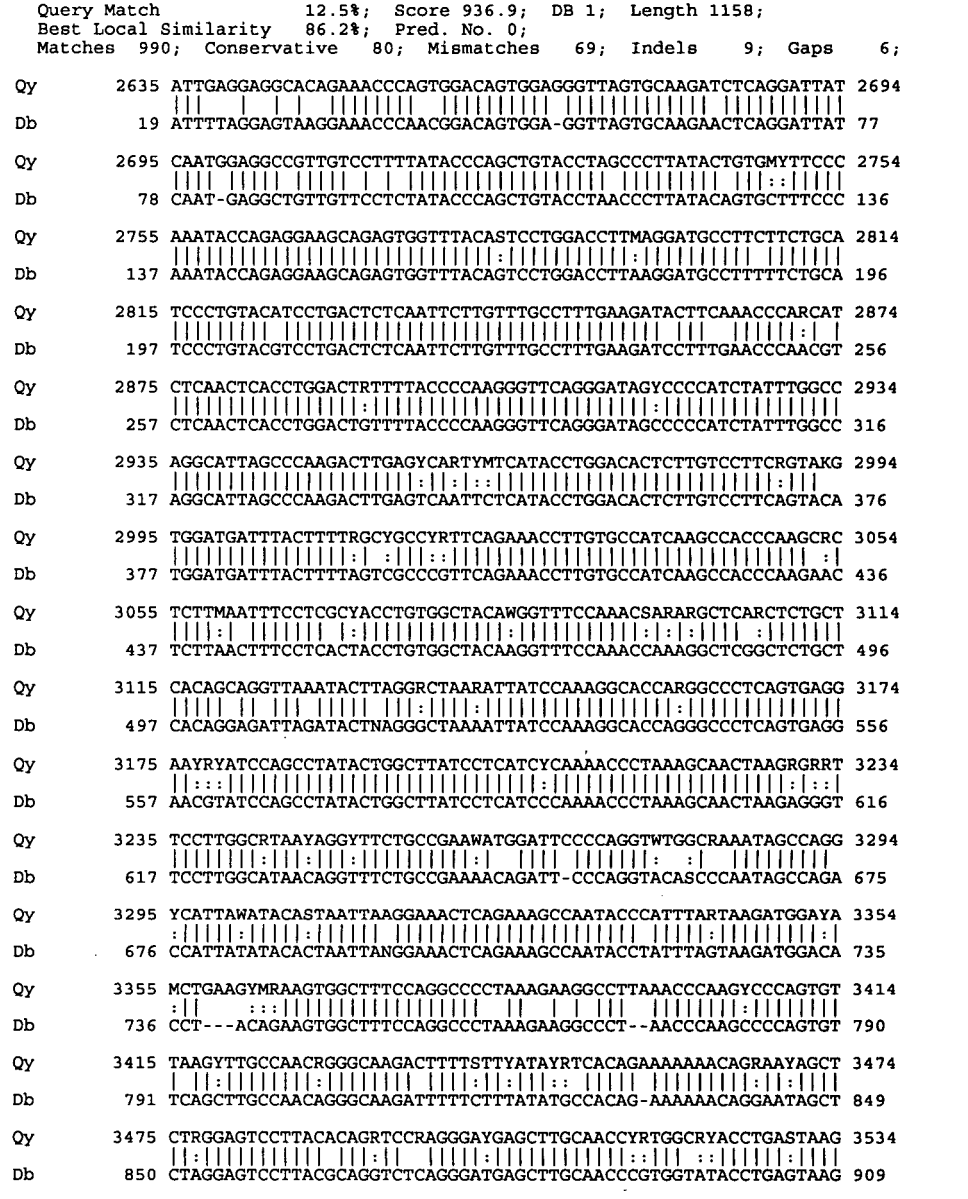
Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316
```



```
RESULT 190
AR036710
LOCUS       AR036710               1158 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5800980.
ACCESSION   AR036710
VERSION     AR036710.1  GI:5954566
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Detection of MSRV1 virus and MSRV2 pathogen and/or infective agent
            associated with multiple sclerosis, by nucleic acid hybridization
  JOURNAL   Patent: US 5800980-A 1 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"
```

Query Match 12.5%; Score 936.9; DB 1; Length 1158;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 990; Conservative 80; Mismatches 69; Indels 9; Gaps 6;



```
Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db       1150 CTTAAACC 1157


RESULT 191
AR076972
LOCUS       AR076972                1158 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 1 from patent US 5962217.
ACCESSION   AR076972
VERSION     AR076972.1  GI:10003718
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and MSRV2 pathogen and/or infective agent associated
            with multiple sclerosis, and biopolymer constituents thereof
  JOURNAL   Patent: US 5962217-A 1 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
Db       1150 CTTAAACC 1157


RESULT 192
AR094693
LOCUS       AR094693                1158 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 1 from patent US 6001987.
ACCESSION   AR094693
VERSION     AR094693.1  GI:10021851
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 1 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"
```

```
  Query Match            12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity  86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                ||| |  | | |||||||| |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                ||||||||| |||||||||||||||||||||||||||||||||| |||  |||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                ||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
                ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
                :|||||:|||||:|||||| |||||||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
                :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
                | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
                ||:|||||||||| |||:||  |||||:|||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
                ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
                ||:: |||||||||||||||||||||||||||||:|||||||| ||||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
                |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
                |||||:|||||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
                ||||| ||
Db       1150 CTTAAACC 1157
```

```
RESULT 193
AR183928
LOCUS       AR183928               1158 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 1 from patent US 6342383.
ACCESSION   AR183928
VERSION     AR183928.1  GI:20227897
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Cell lines and viral isolates associated with multiple sclerosis
  JOURNAL   Patent: US 6342383-A 1 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match            12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity  86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                ||| |  | | |||||||| |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                ||||||||| |||||||||||||||||||||||||||||||||| |||  |||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
                |||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
                |||||||||||||||||||||:||:|::|||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
                ||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
                ||||:| ||||||| |:|||||||||||||:|||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
                ||||| || ||| ||||| |||:||||:|||||||||||||||||:|||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
                ||:::|||||||||||||||||||||||||||:|||||||||||||||||||||:|::|
```

```
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             |||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||| ||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
             ||||| ||
Db       1150 CTTAAACC 1157


RESULT 194
AR343140
LOCUS       AR343140               1158 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 1 from patent US 6579526.
ACCESSION   AR343140
VERSION     AR343140.1  GI:33738642
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 1 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match           12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             |||||||||| |||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
             |||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
             |||||||||||||||||||||:||:|::|||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
             |||||||||||||||||:| :|||::|||||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
             ||||:| ||||||| |:|||||||||||||:|||||||||:|:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
             ||||| || ||| ||||| |||:||||:|||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
             ||:::|||||||||||||||||||||||||||||:|||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|||:||||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||:|||||:|||||| |||||||||||||||||||||| |||||:|||||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:||||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             | ||:|||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:|||||||||| |||:||  |||||:|||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||:: |||||||||||||||||||||||||||||:||||||| ||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149
```

```
Qy       3775 CTTAACCC 3782
              ||||| ||
Db       1150 CTTAAACC 1157


RESULT 195
AR344297
LOCUS       AR344297               1158 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 1 from patent US 6582703.
ACCESSION   AR344297
VERSION     AR344297.1  GI:33740238
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1158)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 1 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1158
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             12.5%;  Score 936.9;  DB 1;  Length 1158;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  990;  Conservative   80;  Mismatches   69;  Indels    9;  Gaps    6;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||   |  | |  ||||||||  |||||||||| |||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              |||| ||||| ||||| | | |||||||||||||||||| |||||||||  |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              ||||||||| |||||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
              ||||||||||||||||:|||||||||||||||||||||||||:|||||||||||||||||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 316

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
              |||||||||||||||||||:||:|::|||||||||||||||||||||||||||:|||
Db        317 AGGCATTAGCCCAAGACTTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTACA 376

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
              |||||||||||||||:| :|||::|||||||||||||||||||||||||||||| :|
Db        377 TGGATGATTTACTTTTAGTCGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGAAC 436

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
              ||||:| ||||||| |:|||||||||||:||||||||||:|:|:|||| :|||||||
Db        437 TCTTAACTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCTGCT 496

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
              ||||| || ||| ||||| |||:||||:||||||||||||||||:|||||||||||||||
Db        497 CACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 556

Qy       3175 AAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRT 3234
              ||:::|||||||||||||||||||||||||:|||||||||||||||||||||:|::|
Db        557 AACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGT 616

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|||:|||||||||:|  |||| |||||||:  :|  |||||||||
Db        617 TCCTTGGCATAACAGGTTTCTGCCGAAAACAGATT-CCCAGGTACASCCCAATAGCCAGA 675

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||:|||||:|||||| ||||||||||||||||||||||| |||||:|||||||:|
Db        676 CCATTATATACACTAATTANGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACA 735

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||    :::|||||||||||||||||||  ||  | | |||  ||||||||:|||||||
Db        736 CCT---ACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGT 790

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
              | ||:|||||||||:|||||||| ||||:||:|||:: ||||| |||||||||:||:||||
Db        791 TCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAG-AAAAAACAGGAATAGCT 849

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:||||||||||| |||:||  |||||:||||||||||||::||| ::|||||:||||
Db        850 CTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 909

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:|||| |||||:||||| ||| ||||||||||
Db        910 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGGGTAATGGNGGCAGTAGCA 969

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||:: |||||||||||||||||||||||||||||:|||||||| ||||||||||||||||
Db        970 GTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTNCTGTGTGGACATCT 1029

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::|||||||||||||||||||||||||||||||||| ||||||||||: ||||
Db       1030 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGTTGTCAGACAACCATTTA 1089

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              ||||| |:|||||||||||||||||: |||||||||| :|||| ||||||||||||||||
Db       1090 CTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGACTGCGCACTTGTGCAACT 1149

Qy       3775 CTTAACCC 3782
              ||||| ||
Db       1150 CTTAAACC 1157


RESULT 196
AY050301
LOCUS       AY050301               957 bp    DNA     linear   PRI 10-OCT-2001
DEFINITION  Homo sapiens clone HERVWSUcl3 envelope protein gene, partial cds.
ACCESSION   AY050301
VERSION     AY050301.1  GI:16024918
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     New HERV-W SU sequences
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2001) Department of Medical Sciences, Section of
            Virology, Dag Hammarkjolds v. 17, Uppsala SE-75185, Sweden
FEATURES             Location/Qualifiers
     source          1. .957
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="HERVWSUcl3"
                     /note="endogenous_virus: HERV-W SU"
     CDS             1. .>957
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="AAL11495.1"
                     /db_xref="GI:16024919"
```

```
                    /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                    GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                    TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                    HTRLVSLFNTALTGLHEVSAQNPTNSWMCLPLHFRPYISTPVPEQWNNFSTEIITTSV
                    LVGPLASNVEIIHTSNLTCVKFSNTIDTTNSQCIRWVTPPTRIVCLPSGIFFVCGTSA
                    YHCLNSFSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVP"

  Query Match             12.2%;  Score 914.6;  DB 1;  Length 957;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  920;  Conservative   14;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||||:||:||||:|||:||||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTC 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
             |||||||||||||||||||||||||||||  |||||||||||||||||||||||||| |||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACAGCCCTCACTGGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
             |||||||||||||| ||||||| |||||||||||| ||||||:|||||| |||||| |
Db        541 CAAAACCCTACTAACAGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAACC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
             |||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db        601 CCTGTACCTGAACAGTGGAACAACTTCAGCACAGAAATAATCACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
             |||||||||| |||||||:||||||||| |||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGCTTCCAATGTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
             ||||| |||| | ||||||||||||||||||||||||||||||||||||||||||||| |
Db        721 AGCAACACTATAGACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
             |||||||||||||| ||||||||||||||||||||||||||| |||||||||| |||||||
Db        781 ATAGTCTGCCTACCTTCAGGAATATTTTTTGTCTGTGGTACTTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
             ||| ||| ||||||||||||||||||||||||||||||||||:|||:||||||||
Db        841 AATAGCTTTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 957


RESULT 197
AY050299
LOCUS       AY050299                957 bp    DNA     linear   PRI 10-OCT-2001
DEFINITION  Homo sapiens clone HERVWSUcl1 envelope protein gene, partial cds.
ACCESSION   AY050299
VERSION     AY050299.1  GI:16024914
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     New HERV-W SU sequences
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2001) Department of Medical Sciences, Section of
            Virology, Dag Hammarkjolds v. 17, Uppsala SE-75185, Sweden
FEATURES             Location/Qualifiers
     source          1. .957
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="HERVWSUcl1"
                     /note="endogenous_virus: HERV-W SU"
     CDS             1. .>957
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="AAL11493.1"
                     /db_xref="GI:16024915"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTALTGLHEVSAQNPTNSWMCLPLHFRPYISTPVPEQWNNFSTEIITTSV
                     LVGPLASNVEIIHTSNLTCVKFSNTIDTTNSQCIRGVTPPTRIVCLPSGIFFVCGTSA
                     YHCLNSFSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVP"

  Query Match             12.2%;  Score 913;  DB 1;  Length 957;
  Best Local Similarity   96.0%;  Pred. No. 0;
  Matches  919;  Conservative   14;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
             ||||:|||||||||||||||||||||||:||:||||:|||:||||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
             ||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTC 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
```

```
Db        361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||||||||||||||||||  |||||||||||||||||||||||||| |||
Db        481 CATACTCGCCTGGTAAGCCTATTTAATACAGCCCTCACTGGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||| ||||||| |||||||||||| ||||||:||||||  |||||| |
Db        541 CAAAACCCTACTAACAGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAACC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              |||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||
Db        601 CCTGTACCTGAACAGTGGAACAACTTCAGCACAGAAATAATCACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||| |||||||:||||||||| ||||||||||||||||||||||||||||||||
Db        661 GGACCTCTTGCTTCCAATGTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||| |||| | |||||||||||||||||||||||| ||||||||||||||||||| |
Db        721 AGCAACACTATAGACACAACCAACTCCCAATGCATCAGGGGGGTAACTCCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||| ||||||||||||||||||||||||| |||||||||| |||||||
Db        781 ATAGTCTGCCTACCTTCAGGAATATTTTTTGTCTGTGGTACTTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||| ||| ||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db        841 AATAGCTTTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 957


RESULT 198
AY050300
LOCUS       AY050300                957 bp    DNA     linear   PRI 10-OCT-2001
DEFINITION  Homo sapiens clone HERVWSUcl2 envelope protein gene, partial cds.
ACCESSION   AY050300
VERSION     AY050300.1  GI:16024916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     New HERV-W SU sequences
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 957)
  AUTHORS   Jern,P., Karlsson,D. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2001) Department of Medical Sciences, Section of
            Virology, Dag Hammarkjolds v. 17, Uppsala SE-75185, Sweden
FEATURES             Location/Qualifiers
     source          1. .957
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="HERVWSUcl2"
                     /note="endogenous virus: HERV-W SU"
     CDS             1. .>957
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="AAL11494.1"
                     /db_xref="GI:16024917"
                     /translation="MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRP
                     GNIDAPSYRSLSKGTPTFTAHTHMPRNCYHSATLCXHANTHYWTGKMINPSCPGGLGV
                     TVCWTYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
                     HTRLVSLFNTALTGLHEVSAQNPTNSWMCLPLHFRPYISTPVPEQWNNFSTEIITTSV
                     LVGPLASNVEIIHTSNLTCVKFSNTIDTTNSQCIRWVTPPTRIVCLPSGIFFVCGTSA
                     YHCLNSFSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVP"

  Query Match             12.2%;  Score 912;  DB 1;  Length 957;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  918;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db         121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         181 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCCNG 240

Qy       5821 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTT 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTC 300

Qy       5881 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTT 360

Qy       5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACAT 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db         361 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 420

Qy       6001 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACC 480

Qy       6061 CATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCC 6120
              |||||||||||||||||||||||||||||  |||||||||||||||||||||||||| |||
Db         481 CATACTCGCCTGGTAAGCCTATTTAATACAGCCCTCACTGGGCTCCATGAGGTCTCAGCC 540

Qy       6121 CAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATC 6180
              |||||||||||||||| ||||||| |||||||||||| ||||||:||||||  |||||| |
Db         541 CAAAACCCTACTAACAGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATACATTTCAACC 600

Qy       6181 CCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 6240
              |||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||
Db         601 CCTGTACCTGAACAGTGGAACAACTTCAGCACAGAAATAATCACCACTTCCGTTTTAGTA 660

Qy       6241 GGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTT 6300
              |||||||||| |||||||:||||||||| ||||||||||||||||||||||||||||||||
Db         661 GGACCTCTTGCTTCCAATGTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTT 720

Qy       6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAA 6360
              ||||| |||| | |||||||||||||||||||||||| ||||||||||||||||||| |
Db         721 AGCAACACTATAGACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGA 780

Qy       6361 ATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG 6420
              |||||||||||||| ||||||||||||||||||||||||| |||||||||| |||||||
Db         781 ATAGTCTGCCTACCTTCAGGAATATTTTTTGTCTGTGGTACTTCAGCCTATCATTGTTTG 840

Qy       6421 AATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTAC 6480
              ||| ||| ||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db         841 AATAGCTTTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTAC 900

Qy       6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 6537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCC 957
```

```
RESULT 199
CQ745034
LOCUS       CQ745034               1080 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30968 from Patent WO02068579.
ACCESSION   CQ745034
VERSION     CQ745034.1  GI:42362241
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 30968 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             12.0%;  Score 897.601;  DB 1;  Length 1080;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches  944;  Conservative   52;  Mismatches   74;  Indels   10;  Gaps    3;

Qy       3548 TGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTG 3607
              |||||||||||||:|:||||||||||:|||||||||  |||||||  |||::|||||||||
Db             2 TGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCAGCAGTAGTAGTCTTAGTATCTG 61

Qy       3608 AAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYR 3667
              |||||||||||  |||||||||:|||||||||||||||||||||||||||||:|||||::
Db            62 AAGCAGTTAAACTAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATG 121

Qy       3668 GCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGG 3727
              |||||||||||||||||||||||||||||||||||||||||||||: ||||||||||:|||||
Db           122 GCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGG 181

Qy       3728 CTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCN 3787
              |||||||||||||:||||||||||||:||||||||||||||||||||||||||||||:|
Db           182 CTCTATTACTTGAAGGGCCAGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCA 241

Qy       3788 CATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAA 3847
              |||||||||||||||| ||||||||||:||:|||||||||| |||:|||||  |||||
Db           242 CATTTCTTCCAGACAATAAAGAAAAGATAGAACATAACTGTCAGCAAGTAATTGTTCAAA 301

Qy       3848 CCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTG 3907
              |||| ||  ||| |||||| | |:||||:|||||:||||||||||:|||| ||| |||
Db           302 CCTACGCTGCTCAAGGGGATCCTTTAGAGGTTCCCTTGACTGATCCTGACC-TCACCTTA 360

Qy       3908 TATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTC 3967
              | ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db           361 TGTACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGTGGGGTATGCAGTGGTC 420

Qy       3968 AGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCA 4027
              |||||||||||||:|||||||||||||||||||||||||||||||| |:||||:|||
Db           421 AGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGTTCAGCTGGCA 480

Qy       4028 GAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATAT 4087
              ||||||||||||:|||:| ||||||||||||||||||||:||||||||:|||||||||||
Db           481 GAACTAATAGCCCTCACTAGGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATAT 540

Qy       4088 ACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
              |||| ||||:|:|||||||:|||||||  |||||||||| |::||||||||:|:|||||||
Db           541 ACAGGCTCTAAGTATGCTTACCTAGTTCTCCATGCCCACGCAGCAATATGGAGAGAAAGG 600

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
              |||||||||||||: ||:||||||||||||:|||||| |||||||||||||| |||||||

Db           601 GAATTCCTAACTTCCAAGGGAACACCTATCAAACATCAAGAAGCCATTAGGATATTATTA 660

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT------CAN 4261
              :||||:||||||||||||:||||||||:||||||||||||:||||||||      |
Db           661 TTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCATAAAAAAAAA 720

Qy       4262 AAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCT 4321
              |||  ||||:|||||||| ||:|||:|||:||||||||||:|| ||||||||:|||||||
Db           721 AAAAAAAAGGAAAGGGAAGTAGAAGGGAACTGCCAAGCAGATGTTGAAGCCAAAAGAGCC 780

Qy       4322 GCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCC 4381
              |||| |||||||||||||||||||||||||||  ||    ||| |||| | |||||| |||
Db           781 ACAAGACAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTGTGGGGTAACCCC 840

Qy       4382 TTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAGG- 4440
              ||| |||||||||||||||||||||||||||| ||||| ||||||| |||||||| ||||
Db           841 CTCCAGGAAACCAAGCCCCAGTACTCAGCAGAAGAAATAGAATGGAGAACCTCAAGAGGA 900

Qy       4441 --CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCA 4498
                | ||| |||||||| ||| || |||||||  ||||||  |||||||||||||||||||
Db           901 CATACTTTCCTCCCCTCAGGATGGCTAGCCACCAAAGAAGAAAAAATACTTTTGCCTGCA 960

Qy       4499 ACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGC 4558
              ||| ||||| |||||||||||||||||| | ||||||||||||||||||| ||||||
Db           961 GCTAACCAATTGAAATTACTTAAAACCCTTTACCAAACCTTTCACTTAGGCATTGATAGC 1020

Qy       4559 ACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGCAR 4618
              | | ||||:||||| |||| ||||||||||||||||||||||||||||||||  |:
Db           1021 AACTATCAGATGGCTAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAATAG 1080


RESULT 200
BD221812/c
LOCUS       BD221812               1216 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses.
ACCESSION   BD221812
VERSION     BD221812.1  GI:33031582
KEYWORDS    JP 2002518051-A/7.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1216)
  AUTHORS   Alliel,P.M., Perin,J.P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: JP 2002518051-A 7 25-JUN-2002;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002518051-A/7
            PD   25-JUN-2002
            PF   23-JUN-1999 JP 2000556036
            PR   23-JUN-1998 FR     98/07920
            PI   PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC
            C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
            A61P25/00,
            PC   A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC   Nucleic
            sequence and deduced protein sequence family with  CC
            human endogenous
            CC   retroviral motifs, and their uses
            FH   Key             Location/Qualifiers
            FT   source          1. .1216
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1216
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             11.9%;  Score 894.401;  DB 1;  Length 1216;
  Best Local Similarity   84.6%;  Pred. No. 0;
```

```
  Matches 1003;  Conservative   21;  Mismatches  102;  Indels   60;  Gaps    3;

Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATT 4729
             |||||||||||||||:||||||:|||||||||||||| | |:|:||:|||||||||| ||||
Db       1216 AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACTCAGGAGAAGACTGGCAACTAGATT 1157

Qy       4730 TTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCA 4789
             |||||||| ||||||| ||||||||||||||||||||||||||||||:||| |||||
Db       1156 TTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATATTTTCA 1097

Qy       4790 CGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCAC 4849
             | |||||||||:|||||||||  |||||||||||||||| |||||| ||||||||||||||
Db       1096 CTGGTTGGGCAGAGGCCTTCCTTTGTAGGACAGAAAAGGACCAAGATGTAATAAAGGCAC 1037

Qy       4850 TAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCC 4909
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1036 TAATTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATGGCC 977

Qy       4910 CTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACT 4969
             | | |||| |||| |||||||||||||||||||||||||||||| ||||| |||| ||||||||
Db        976 CCGTTTTCAAGGCTACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCACT 917

Qy       4970 TACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCA 5029
             |||||||||| || |||||||| |||||||| || ||| |||||||||||||| :|||||
Db        916 TACACTGCGCGTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAATGAAGCACTCA 857

Qy       5030 AAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTA 5089
             || ||||| ||||||||| |||||| |||||||||||  |||||||||:||||||||||||
Db        856 AACGACATTTAAAAAAGCTAACCCAAGAAACCCACCTTGCATGGCCTGCTCTGTTGCCTA 797

Qy       5090 TAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGC 5149
             |||||||  |||||||   |||| |||||||||||||||||||||||||||||||||||||
Db        796 CAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAAGCAGGACTTAGCCCATACAAAATGC 737

Qy       5150 TGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAG 5208
             |||||||| |||||||| ||||||| |||||||||||||||| | |||  |||| |||||
Db        736 TGTATGGATGGCCCTTCCTAACCAAGGACCTTGTGCTTGACCGAGAGATGGCCAGCTTAG 677

Qy       5209 TTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCT 5268
             |||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||
Db        676 TTGCAGACATCACCTCCTTGGCCAAATATCAACAAGTTCTTAAAACATTACAGGGAACCT 617

Qy       5269 ATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCC 5328
             | || |||||||||||||  || |||| |||||: ||||||||||||||||||||||||||
Db        616 GTTCCCGAGAAGAGGGAAAGTAATTATTGCACCCTGGTGACATGGTATTAGTCAAGTCCC 557

Qy       5329 TTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 5388
             |||: ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        556 TTCCTTCTAAGTCCCCATCCTTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 497

Qy       5389 YTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 5448
             :|||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        496 CTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGAATACATCACACTCGAGTCA 437

Qy       5449 AATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 5508
             ||  ||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        436 AACTCTGTATACTGCCAAAGGAACCCGAAAATCCAGGAGACAACGCTAGCTATTCCTGTG 377

Qy       5509 AACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAAT 5568
             |||||||||||||| ||| |||||||||||| | ||||||  |||||||||||||| |||
Db        376 AACCTCTAGAGGATCTGCACCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAAT 317

Qy       5569 CATAAATCCCCCATG---------------------------------------------- 5583
             | || |||| |   |
Db        316 CGTAGATCCTCATGGCCCTCCTTGTCATATTTTTTTCTTTACTGTTCTCTTACCCCCTTT 257

Qy       5584 ------------GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTT 5630
                         :|||||||||||||||||||||||:||:|||||:|||:||||||:|||
Db        256 CACTCTCACTGAACCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTT 197

Qy       5631 CACTCTCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTT 5690
             |||||||||||||||||||||||||| |||||  |||||||||||||| |:|||:|||| |
Db        196 CACTCTCACTGCACCCCCTCCATGCCACTGTACTACCAGTAGCTCCCATTACCAAGAGCT 137
```

```
Qy       5691 TCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAA 5750
             |||||||||||||| || ||| |||||||||||||| |||||||||||||||| |||:|||
Db        136 TCTATGGAGAATGCGGCTTCCTGGAAATATTGATGACCCATCGTATAGGAGTTTTTCTAA 77

Qy       5751 GGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||| |||||  |||||  |||||||| ||||| |||   ||| |
Db         76 AGGAAACCCCATTTTCACCACCCACACCTATATGACCCTGCACTTC 31


RESULT 201
AR699761/c
LOCUS       AR699761               1216 bp.   DNA     linear   PAT 14-SEP-2005
DEFINITION  Sequence 7 from patent US 6919438.
ACCESSION   AR699761
VERSION     AR699761.1  GI:75205551
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1216)
  AUTHORS   Alliel,P.M., Perin,J.-P. and Rieger,F.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: US 6919438-A 7 19-JUL-2005;
            Institut National de la Sante et de la Recherche Medical-Inserm;
            Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1216
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             11.9%;  Score 894.401;  DB 1;  Length 1216;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches 1003;  Conservative   21;  Mismatches  102;  Indels   60;  Gaps    3;

Qy       4671 AAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT-GATT 4729
             |||||||||||||:||||||:|||||||||||||| | |:|:||:|||||||||| ||||
Db       1216 AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACTCAGGAGAAGACTGGCAACTAGATT 1157

Qy       4730 TTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACTTTCA 4789
             |||||||| ||||||| ||||||||||||||||||||||||||||||:||| |||||
Db       1156 TTACCCACATGCCCAAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATATTTTCA 1097

Qy       4790 CGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCAC 4849
             | |||||||||:|||||||||  |||||||||||||||| |||||| ||||||||||||||
Db       1096 CTGGTTGGGCAGAGGCCTTCCTTTGTAGGACAGAAAAGGACCAAGATGTAATAAAGGCAC 1037

Qy       4850 TAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCC 4909
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1036 TAATTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATGGCC 977

Qy       4910 CTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACT 4969
             | | |||| |||| |||||||||||||||||||||||||||||| ||||| |||| ||||||||
Db        976 CCGTTTTCAAGGCTACAGTAACCCAGGGAGTATCCCAGGCATTAGGCATACAATATCACT 917

Qy       4970 TACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCA 5029
             |||||||||| || |||||||| |||||||| || ||| |||||||||||||| :|||||
Db        916 TACACTGCGCGTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAATGAAGCACTCA 857

Qy       5030 AAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCTA 5089
             || ||||| ||||||||| |||||| |||||||||||  |||||||||:||||||||||||
Db        856 AACGACATTTAAAAAAGCTAACCCAAGAAACCCACCTTGCATGGCCTGCTCTGTTGCCTA 797

Qy       5090 TAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGC 5149
             |||||||  |||||||   |||| |||||||||||||||||||||||||||||||||||||
Db        796 CAGCCTTACTAAGAATCCAAAACTCTCCCCAAAAAGCAGGACTTAGCCCATACAAAATGC 737

Qy       5150 TGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAACTTAG 5208
             |||||||| |||||||| ||||||| |||||||||||||||| | |||  |||| |||||
Db        736 TGTATGGATGGCCCTTCCTAACCAAGGACCTTGTGCTTGACCGAGAGATGGCCAGCTTAG 677
```



```
RESULT 202
AX007984/c
LOCUS       AX007984               1216 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 7 from Patent WO9967395.
ACCESSION   AX007984
VERSION     AX007984.1  GI:9995681
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Perin,J.P., Rieger,F. and Alliel,P.M.
  TITLE     Nucleic sequence and deduced protein sequence family with human
            endogenous retroviral motifs, and their uses
  JOURNAL   Patent: WO 9967395-A 7 29-DEC-1999;
            INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
            FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES             Location/Qualifiers
     source          1. .1216
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             11.9%;  Score 894.401;  DB 1;  Length 1216;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches 1003;  Conservative   21;  Mismatches  102;  Indels   60;  Gaps    3;
```



```
Qy       5691 TCTATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAA 5750
             |||||||||||||| || ||| |||||||||||| |||||||||||||||| |||:|||
Db        136 TCTATGGAGAATGCGGCTTCCTGGAAATATTGATGACCCATCGTATAGGAGTTTTTCTAA 77

Qy       5751 GGGAACCCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||| |||||  |||||  |||||||| ||||| |||   ||| |
Db         76 AGGAAACCCCATTTTCACCACCCACACCTATATGACCCTGCACTTC 31


RESULT 203
CQ746663
LOCUS       CQ746663                 918 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 32597 from Patent WO02068579.
ACCESSION   CQ746663
VERSION     CQ746663.1  GI:42368093
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 32597 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             10.8%;  Score 810.2;  DB 1;  Length 918;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  822;  Conservative   55;  Mismatches   41;  Indels    4;  Gaps    2;

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::||||||:||||||||:|||||||||||||||||||||:|:||||||||||: |
Db          1 ATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACAG 60

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||||||| |||||||||||::|||||||||||||||||||| ||||||||| :||||||
Db         61 GTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGAAAGAGAT 120

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              |||||||||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db        121 CTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGG 180

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              |||||||||||:|||||||||||:|||||||||| | ||||||:| |||||||||||:|
Db        181 CTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCAGTGCTGCGA 240

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATA 3816
              |||||||||||||||||||||||||:| |||||||||||| ||||||||||||||||||
Db        241 CTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGGCAATGAAGAAAAGATA 300

Qy       3817 RAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGAR 3876
              :||:|||||||||||||:|||||  |||||||| ||||||| ||||||||||:||||:
Db        301 GAACATAACTGTCAACAAGTAATTGTTCAAACCTACGCCACTCAAGGGGACCTTTTAGAG 360

Qy       3877 GTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 3936
              |||||: |||||||||:|||| |||||||||||||||||||||||||| ||||||||
Db        361 GTTCCCCTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCTTCGTAGAAAA 419

Qy       3937 AGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCC 3996
              ||| |||| |||||:||||||||||| ||||||||||||||||:||||||||||||||
Db        420 AGGGCTTCAAAAAGTGGGGTATGCAGTTGTCAGTGATAATGGAATACTTGAAAGTAATCC 479

Qy       3997 CCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGA 4056
              |||||||||||||||||||:||||:||||||||||||||:|||:|  |||| ||||
Db        480 CCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCATTCAGGCAGTAGA 539

Qy       4057 ATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNT 4116
              |||   |||||:||||||||:|||||||||||||||||||:|:||||||:||||||| |
Db        540 ATT---AGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTACCTAGTCCT 596

Qy       4117 CCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTAT 4176
              ||||||||||::||||||||:|:|||||||| ||||||||||||:|||:||||||||||
Db        597 CCATGCCCATGCAGCAATATGGAGAGAAAGGGGATTCCTAACTTCCGAGGGAACACCTAT 656

Qy       4177 CAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGM 4236
              ||:|||||||||||||||||||:||||||||:|||| |||||||||||||:||||||||:
Db        657 CAAACATCAGGAAGCCATTAGGAGATTATTATTGGCGGTACAGAAACCTAAAGAGGTGGC 716

Qy       4237 AGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCA 4296
              |||||||||||||:|||||||||| |||||||||:||||||||||:|||:|||:|||||
Db        717 AGTCTTACACTGCTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACTGCCA 776

Qy       4297 AGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAA 4356
              ||| :||||||||||:|||||||||||||||||||||||||||||||||||||||||| |
Db        777 AGCGGATATTGAAGCAAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATAGA 836

Qy       4357 ACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGA 4416
              |   ||| ||||||  |||||||||| || |||||||||||||||||||||||| || |||
Db        837 AGGACCCCTAGTATGTGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTCAAAAGAAGA 896

Qy       4417 AACAGAATGGGGAACCTCACGA 4438
              || |||||||||||||||| ||
Db        897 AATAGAATGGGGAACCTCATGA 918


RESULT 204
CQ747448
LOCUS       CQ747448                1038 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33382 from Patent WO02068579.
ACCESSION   CQ747448
VERSION     CQ747448.1  GI:42370534
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 33382 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1038
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             10.5%;  Score 787.1;  DB 1;  Length 1038;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  884;  Conservative   54;  Mismatches   83;  Indels   19;  Gaps    7;

Qy       2424 CTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAA 2483
              |||||||||||||||||||||||||| |||||||| ||| |||||||| || | |||||
Db          11 CTGAAGTCTGGGCAACAGAAGGACAATGTGGACGAG-CAATGAATGCCCATCGTATTCAA 69

Qy       2484 GTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAG 2543
              |||||||||||||||| | |||||||||||||||||||||||||| ||||    ||
Db          70 ATTAAACTAAAGGATTCCGCCTCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACATGAG 129

Qy       2544 GCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCA 2603
              ||||||||| || |||||||||   |||||||| |||||||||||||| |||||||| ||
Db         130 GCCCAACAAAGACTCCAAAAGA---TTAAGGACCTAAAAGCCCAAGGCCTAGTAAAATCA 186

Qy       2604 TGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACA 2663
              ||||  |  ||||||| ||   |        |||   |  | |||||||||| ||||||
```

```
Db        187 TGCAGTAGCCCCTGCAATACCCC-------AATTTTAGGAGTACAGAAACCCAATGGACA 239

Qy       2664 GTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCA 2723
             ||||| ||||||||||||||||  |  || |  |    |||||||| ||| | ||||||||
Db        240 GTGGA-GGTTAGTGCAAGATCGGATTATCAGTA----AGGCCGTTTTCCCTCTATACCCA 294

Qy       2724 GCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACA 2783
             |||||||||| |||||||||| ||::||| ||||||||||||||||||||||||||||||
Db        295 GCTGTACCTAAACCTTATACTCTGCTTTCTCAAATACCAGAGGAAGCAGAGTGGTTTACA 354

Qy       2784 STCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTG 2843
             :||||| ||||||:|||||||||| |||||||| ||||||||||||||||||||||||||
Db        355 GTCCTGAACCTTAAGGATGCCTTTTTCTGCATCTCTGTACATCCTGACTCTCAATTCTTT 414

Qy       2844 TTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAA 2903
             |||||||||||||| ||| ||||||:  |||||||||||||||||:||||||| |||
Db        415 TTTGCCTTTGAAGATCCTTTGAACCCAATGTCTCAACTCACCTGGACTGTTTTACCGCAA 474

Qy       2904 GGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYM 2963
             |||||||||||||:|||||||||||||||| ||||||||||||||||:||:|:::
Db        475 GGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGTATTAGCCCAAGACTTGAGCCAGTTC 534

Qy       2964 TCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCA 3023
             |||||||||||||||||||||||:|||:|||||||||||| ||||:||: ||::||||
Db        535 TCATACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTAATTTTAGCCACCCATTCA 594

Qy       3024 GAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTAC 3083
             |||||||||||||||||||||||||||:|||||:|||||||| ||:||||||||||||
Db        595 GAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTAGCCACCTGTGGCTAC 654

Qy       3084 AWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARA 3143
             |:|||||||||||:| |:|||||:|||||||||||||||||||||||||||| :||||:|
Db        655 AAGGTTTCCAAACCACAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGAGCTAAAA 714

Qy       3144 TTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCT 3203
             |||||||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||
Db        715 TTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCT 774

Qy       3204 CATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWA 3263
              |||:|||||||||||||||||||| :| :||||||||| |||:|||:|||||||||:|
Db        775 TATCCCAAAACCCTAAAGCAACTAACAGCATTCCTTGGCTTAACAGGTTTCTGCCGAATA 834

Qy       3264 TGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCA 3323
             |||||| ||||| |: |||:|||||||||| :|||||:|||| :|||||||||||||||
Db        835 TGGATT-CCCAGATACGGCGAAATAGCCAGACCATTATATACGCTAATTAAGGAAACTCA 893

Qy       3324 GAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCC 3383
             |||||||||||||||||||:|||||||||:|:|||||||:::|||  |||| ||||||||
Db        894 GAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCAGCTTCCCAGGCCCT 953

Qy       3384 TAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTST 3443
              ||  | | |||  ||||||||:||||||||||||:||||||| :|||||||||||||:|
Db        954 AAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAATGGGGCAAGACTTTTCT 1011

Qy       3444 TYATAYRTCACAGAAAAAAA 3463
              |:|||::|||| |||| |
Db       1012 TTATATGTCACACAAAAACA 1031


RESULT 205
CQ746367
LOCUS       CQ746367                 918 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 32301 from Patent WO02068579.
ACCESSION   CQ746367
VERSION     CQ746367.1  GI:42367011
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 32301 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             10.0%;  Score 747.7;  DB 1;  Length 918;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  820;  Conservative   31;  Mismatches   59;  Indels   11;  Gaps    5;

Qy       2299 ATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATT 2358
             |||| |||||||| |||| ||||||||| |||||| ||||||| ||||||||||||||||
Db          1 ATGCCTGAAAGCCTCACTCCCTTGTTAGAGAGAGAGATTCTAGTAAAAGCAGGGGCCATT 60

Qy       2359 ATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAAT 2418
             |||||||||||||||||||||||||||||| ||||| || | ||||||||||||||||||
Db         61 ATACACCTGAACATAGGAGAAGGAACACCCATTTGTCGT-CTCCTGCTTGAGGAAGGAAT 119

Qy       2419 TAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTG 2478
             || |||||||||||||||||||||||||||||||||| || |||||||||||| |  ||
Db        120 TAGTCCTGAAGTCTGGGCAACAGAAGGACAATATGGACAAG-CAAAGAATGCCCATTTTG 178

Qy       2479 TTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGAC 2538
             | |||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db        179 TCCAAGTTAAACTAAAGGATTCCACCTCCTTTCCCTATCAAAGGCAGTACCCCCTCAGAC 238

Qy       2539 CCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAA 2598
             || |||||||||| || |||||||||||||||||| |||||||||||||| ||||||
Db        239 CCGAGGCCCAACAAAGACTCCAAAAGATTGTTAAGGACCTAAAAGCCCAAGGCCTAGTAA 298

Qy       2599 AACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGT 2658
             |||||| ||  |  ||||||| || | |     |||   |  | ||||||||||
Db        299 AACCATACAGTAGCCCCTGCAATACTCC-------AATTTTAGGAGTACAGAAACCCAAA 351

Qy       2659 GGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTAT 2718
             |||||||| |||||||||||||||||||||||||||||| ||||| ||| | | | |||
Db        352 AGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT-GAGGCTGTTTTTCCTCTAT 409

Qy       2719 ACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGT 2778
             |||||||| |||||||||||||||| ||::|||||||||||||||||||||||||||||
Db        410 ACCCAGCTATACCTAGCCCTTATACTCTGCTTTCCCAAATACCAGAGGAAGCAGAGTGGT 469

Qy       2779 TTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAAT 2838
             |||||:|||||||| |||:|||||||||| ||||||||||||||||||||||||||||||
Db        470 TTACAGTCCTGGATCTTAAGGATGCCTTTTTCTGCATCCCTGTACATCCTGACTCTCAAT 529

Qy       2839 TCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTAC 2898
             |||||||||| |||||||||| ||||||||||:|||||||||||||||||||||:||||||
Db        530 TCTTGTTTGCATTTGAAGATCCTTCAAACCCAACATCTCAACTCACCTGGACTGTTTTAC 589

Qy       2899 CCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYC 2958
             |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:|
Db        590 CCCAAGGGTTCAGGGATAGTCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCC 649

Qy       2959 ARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCY 3018
             |:|::|||||||||||||||| ||||||||:||   |||||||||||||||:||:|||:
Db        650 AGTTCTCATACCTGGACACTCCTGTCCTTCAGTGCATGGATGATTTACTTTTAGCTGCCC 709

Qy       3019 RTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTG 3078
             :|||||||||||||||||||||||||||||||||:|||||:|||||||||| |:|||||||
Db        710 GTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTCACTACCTGTG 769

Qy       3079 GCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRC 3138
             ||||||:|||||||||||:|:|:|:|||||:|||||||||||||||||||| ||||||:|
Db        770 GCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATGCTTAGGGC 829

Qy       3139 TAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTT 3198
             |||:|||||||||| |||||:|||||||||||||||:::|||||||||||||||||||||
```

```
Db        830 TAAAATTATCCAAAGTCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTT 889

Qy       3199 ATCCTCATCYCAAAACCCTAA 3219
              ||||| |||:||||||||||||
Db        890 ATCCTTATCCCAAAACCCTAA 910


RESULT 206
CQ749099
LOCUS       CQ749099                993 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 35033 from Patent WO02068579.
ACCESSION   CQ749099
VERSION     CQ749099.1  GI:42376854
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 35033 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .993
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             9.9%;  Score 741;  DB 1;  Length 993;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  875;  Conservative   13;  Mismatches   94;  Indels   12;  Gaps    8;

Qy       1049 GAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAA 1108
              ||||| | ||||||| |||||||||||||| |:||||:: ||| |||||||| | | |||| |
Db         11 GAGAACTTGGCCCAGTCAGAGTGCATGTACCTTTTTCCCTCTCAGACTTGAGGAAAATTA 70

Qy       1109 AAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTACAAGGGT 1168
              ||| |||| ||||||||||||||||||:||||||||||||||||:|||||||||||||||
Db         71 AAATAGACCTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGT 130

Qy       1169 TAGGACAATTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTAAATCAGACACTA 1228
              || |||||| |||||||||||||||||||||||| |||| ||||||| ||||| |||||
Db        131 TAAGACAATCCTTTGATCTGACATGGAGAGATATA-ATGTTACTGCTAGATCAGTCACTA 189

Qy       1229 ACCCCAAATGAGAGAAGTGCCACCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTAT 1288
              |||||||||||| ||||||| ||| |||||||||| |||:||||||||||||||||||||
Db        190 ACCCCAAATGAGAAAAGTGCCGCCAAAACTGCAGCCCGAGAGTTTGGCGATCTCTGGTAT 249

Qy       1289 CTCAGTCAGGTCAATGGATANGGATGACAACAGAAGGAAAGANAATGATTCCCCACAGGC 1348
              |||||||||||| | |   | |||||||||| |||||||| ||  |||||| ||||||
Db        250 CTCAGTCAGGTCGAGGATAA--GATGACAACA-CAGGAAAGAGAACAATTCCCTACAGGC 306

Qy       1349 CAGCARGCAGTTCCCAGTCTASACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGA 1408
              |||||:|||||||||||||| ||:||||||| ||||      | |       |    |||||
Db        307 CAGCAGGCAGTTCCCAGTGTAGACCCTCACTGGG----ATGCAGAATCAGAACACGGAGA 362

Qy       1409 TTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAA 1468
              |||||||  |||||||||||||||||||  |||||:|||||||||| |||||||:|||| |
Db        363 TTGGTGCCACAGACATTTGCTAACTTGCATGCTAGAAGGACTAAAGAAAACTAGGAAG-A 421

Qy       1469 ARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACT 1528
              |: |||:|||||||||||||||||||| |||||||  |||| | | ||||||||||||
Db        422 AGACTATGAATTACTCAATGATGTCCACTATAACACA-AGGAAAGAATGAAAATCCTACT 480

Qy       1529 GCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCT 1588
              |||||||||||||||||||||||| |||||||||||| | ||||||||||||||||||||
Db        481 GCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTGTCACCTGACTCT 540

Qy       1589 TCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG- 1647
              |||||||||||||||||||||||  ||| ||||||| |||||||| || |||||||||
Db        541 ATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTTGTAGACATTAGA 600

Qy       1648 AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTG 1707
              |||||||||||||||||||||| ||||||| |||||||||||||||||||||||||||||
Db        601 AAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACCCTATTGAACTTG 660

Qy       1708 GCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGAT 1767
              |||||:|||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        661 GCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA- 719

Qy       1768 TAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGG 1827
              |||||| |||||||||||||||||||||| || |||||||||| |||||||||| ||||||
Db        720 TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGG 779

Qy       1828 AAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACA 1887
              || ||||||| |||||||||||  |||||||||||||||||||||||||||||||| || | ||
Db        780 AACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCA 839

Qy       1888 AGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTT 1947
              |||||||||||||||||||||  | ||||| |||||||||| ||||||||||||||||||
Db        840 AGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTT 899

Qy       1948 ATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGA 2007
              || |||||||||||||||||| |||||||||||||||||||||||| | |||||||||
Db        900 ATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGA 959

Qy       2008 AGCCACTAACCAGATGATCCAGCAGCAGGACTGA 2041
              |||||||||||||| ||||||||||||||||||||
Db        960 AGCCACTAACCAGAAGATCCAGCAGCAGGACTGA 993


RESULT 207
BV465687
LOCUS       BV465687                892 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P64967RA6.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV465687
VERSION     BV465687.1  GI:62225115
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 892
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
```

chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .892
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>892

Query Match 9.9%; Score 740.4; DB 1; Length 892;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 812; Conservative 7; Mismatches 73; Indels 4; Gaps 3;

```
Qy       5928 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACT 5987
             ||||||||||||||||||||||||||||||||||          || | | | |||||||
Db             1 TGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAACA--CCCTCAACTGACCCCCCAACT 58

Qy       5988 CACCSGGGTACATGGCACCTCTAGCCCCTACAAAGGACTAGATCTCTCAAAACTACATGA 6047
              ||| |||||||  | ||| |   | || | |||||||||||| |||||||||| ||| ||
Db            59 GACCTGGGTACAAAGTACCCCCCCCTCCAAGAAAGGACTAGACCTCTCAAAACCACAGGA 118

Qy       6048 AACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCA 6107
             |||||||| |||||||||||| ||||||||| |||||||||| ||||||||| |||||||
Db         119 AACCCTCCATACCCATACTCGGCTGGTAAGCTTATTTAATACTACCCTCACTAGGCTCCA 178

Qy       6108 TGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCC 6167
             | |||| |||||||||||||||||||||||||||| |||||||||||| ||||||:|||
Db         179 TAAGGTTTCGGCCCAAAACCCTACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCC 238

Qy       6168 ATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCAC 6227
             |||   |||||||||||||||||||||||||||||||| ||||||| ||||||||||||||
Db         239 ATA-CATTCAATCCCTGTACCTGAACAATGGAACAACTCCAGCACAAAAATAAACACCAC 297

Qy       6228 TTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCAC 6287
             ||||||||||||||||||||||||||||||:|||| |||||||||||||||||||| ||
Db         298 TTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAGATAACCCATACCTCAAACCTTAC 357

Qy       6288 CTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAAC 6347
             |||||||||||||| |||||||| | ||||||||||||||||||||||||||||||||
Db         358 CTGTGTAAAATTTAACAATACTATAGACACAACCAACTCCCAATGCATCAGGTGGGTAAC 417

Qy       6348 TCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGC 6407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         418 TCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGA 477

Qy       6408 CTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCC 6467
             |||||||||||||||||| |||||||||||||||| |||||||||||||| ||  |||
Db         478 CTATCGTTGTTTGAATGGCCCTTCAGAATCTATGTGCGTCCTCTCATTCTTACTG-TCCC 536

Qy       6468 YATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAA 6527
             :|||:|||| |||||||||||||||||||||| |||||||| | ||| |||| |||||||
```

```
Db         537 TATGACCATTTACACTGAACAAGATTTATACAATTATGTCACACCTACGCCCGGCAACAA 596

Qy       6528 AAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGG 6587
             |||||||||||||||||||||||| | ||||||||||||||||    ||||||||||
Db         597 AAGAGTACCCATTCTTCCTTTTGTTGTTGGAGCAGGAGTGCTAGGCAGACTAGGTACTGA 656

Qy       6588 CATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGG 6647
             |||||  || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         657 CATTGATGGCATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGG 716

Qy       6648 GGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGC 6707
              |||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||
Db         717 TGACATGGAACAGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAATTTAACTCCCTAGC 776

Qy       6708 AGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAAC 6767
             |||||| |||||||:|||||||||||||||||||:||||||| | ||:||||||| |||
Db         777 AGCAGTGGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCTCCGAAAGAGGGGAAAC 836

Qy       6768 CTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCA 6823
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         837 CTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGAAATCGTCA 892
```

RESULT 208
BV619299
LOCUS BV619299 870 bp DNA linear STS 15-APR-2005
DEFINITION S215P61160RC6.T0 Clara Pan troglodytes troglodytes STS genomic, sequence tagged site.
ACCESSION BV619299
VERSION BV619299.1 GI:62609477
KEYWORDS STS.
SOURCE Pan troglodytes troglodytes
ORGANISM Pan troglodytes troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 870)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 870
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely

placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.

```
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Clara"
     STS             <1. .>870
```

Query Match 9.6%; Score 715.301; DB 1; Length 870;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 787; Conservative 6; Mismatches 73; Indels 3; Gaps 3;

```
Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
Db          3 ATTCCACTGTACGACCAGTAGCTCCCCTTACC-AGCGTTTCTATGGAGACCTCTGCGTCC 61

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
Db         62 CAGACATATTGATGCCCCATCGTATAGGAGTTTATCTAAGGGAAACCCCGCCTTCACCGC 121

Qy       5772 CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATAC 5831
Db        122 CCAGACCCATATGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATAC 181

Qy       5832 TCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGT 5891
Db        182 TCATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGCCACTGT 241

Qy       5892 CTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC 5951
Db        242 CTGTTGGACTTAGTTCACCCATACTGGTATGTCTGATAGGGGTGGAGTTCAAGATCAGGC 301

Qy       5952 AAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAG 6011
Db        302 AAGAGAAAAACAAGTAAAGGAAGTAATCTCCCAACTGACCTGAGTACATAGCACGCCTAG 361

Qy       6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCT 6071
Db        362 CTCCTACAAAGGACTAGATCTCTCAAAACCACATGAAACCCTCCGTACCCATACTCGCCT 421

Qy       6072 GGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTAC 6131
Db        422 GGTAAGCCTATTTAATACCACCCTCACTGGGTTTCATGAGGTCTCGGCCCAAAACCCTAC 481

Qy       6132 TAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGA 6191
Db        482 TAACTGTTGGATGTGCCTCCCTCTGCACTTCAGGCCATACATTTCAATCCCTGTACCTGA 541

Qy       6192 ACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 6251
Db        542 ACAAGGGAACAACTTCAACACATAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGT 601

Qy       6252 TTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTAC 6311
Db        602 TTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGAT 661

Qy       6312 ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCT 6371
Db        662 AGACACAACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCT 720

Qy       6372 ACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTC 6431
Db         721 ACCCTCAGGAATATTTTTTGTCTGTGGTAACTCAGCCTATATGTGTTTGAAAATGCTCTC 780

Qy       6432 AGAATCTATGTGCTTCCTCTCATTCTTAGTG-CCCCCYATGRCCATCTACACTGAACAAG 6490
Db         781 AGAATCTATGTGCTTCCTGTCCTATCTAGTGCCCCCCCGTGACCATCTACACTGAAACAG 840

Qy       6491 ATTTATACAGTTATGTCATATCTAAGCCC 6519
Db         841 ATTCATACTATCATGTCATACCTAACCCC 869
```

RESULT 209
CQ058583
LOCUS CQ058583 792 bp DNA linear PAT 19-JAN-2004
DEFINITION Sequence 9403 from Patent WO0157270.
ACCESSION CQ058583
VERSION CQ058583.1 GI:41033089
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human breast and hbl 100 cells
JOURNAL Patent: WO 0157270-A 9403 09-AUG-2001;
Aeomica, Inc. (US)

```
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN HBL100, SIGNAL = 6
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"
```

Query Match 9.5%; Score 712.9; DB 1; Length 792;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 746; Conservative 5; Mismatches 41; Indels 3; Gaps 1;

```
Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
```

```
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             ||||  |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             |||||  | ||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||||  | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             ||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 210
CQ077936
LOCUS       CQ077936                792 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 13736 from Patent WO0157278.
ACCESSION   CQ077936
VERSION     CQ077936.1  GI:41047805
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 13736 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN HELA, SIGNAL = 7.2
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||||||| |||||||||| ||||  ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             ||||  |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             |||||  | ||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||||  | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             ||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 211
CQ108942
LOCUS       CQ108942                792 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17801 from Patent WO0157272.
ACCESSION   CQ108942
VERSION     CQ108942.1  GI:41077995
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 17801 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
```

EXPRESSED IN PLACENTA, SIGNAL = 20
NT HIT: AF208161.1, EVALUE 0.00e+00
EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
SWISSPROT HIT: P10269, EVALUE 2.00e-16"

Query Match 9.5%; Score 712.9; DB 1; Length 792;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 746; Conservative 5; Mismatches 41; Indels 3; Gaps 1;

```
Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:|||||||
Db        778 TTTAGACTTGCTAAC 792
```

RESULT 212
CQ147575
LOCUS CQ147575 792 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 17597 from Patent WO0157276.
ACCESSION CQ147575
VERSION CQ147575.1 GI:41154925
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 17597 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .792
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002346.1
EXPRESSED IN BONE MARROW, SIGNAL = 6.3
NT HIT: AF208161.1, EVALUE 0.00e+00
EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
SWISSPROT HIT: P10269, EVALUE 2.00e-16"

Query Match 9.5%; Score 712.9; DB 1; Length 792;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 746; Conservative 5; Mismatches 41; Indels 3; Gaps 1;

```
Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||||| ||||||||| |||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717
```

```
Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 213
CQ182942
LOCUS       CQ182942                792 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14338 from Patent WO0157274.
ACCESSION   CQ182942
VERSION     CQ182942.1  GI:41177728
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 14338 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN HEART, SIGNAL = 5.4
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||| ||| ||||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||| ||||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 214
CQ207358
LOCUS       CQ207358                792 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9635 from Patent WO0157271.
ACCESSION   CQ207358
VERSION     CQ207358.1  GI:41193527
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 9635 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN BT474, SIGNAL = 5.9
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||| ||| ||||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
```

```
                ||||||||| |||||||||||| |||| |:|||||| |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
                |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
                ||||:||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
                |||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
                |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
                ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
                |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
                |||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
                |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
                ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 215
CQ230826
LOCUS       CQ230826                792 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17665 from Patent WO0157273.
ACCESSION   CQ230826
VERSION     CQ230826.1  GI:41214044
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 17665 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
                |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
                |||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
                |||||||||  |||||||||||  |||| |:|||||  |||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 6254
                |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
                ||||:||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
                |||| |||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
                |||| |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
                ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
                |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
                |||||||||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
                |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
                ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 216
CQ268963
LOCUS       CQ268963                792 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 17224 from Patent WO0157277.
ACCESSION   CQ268963
```

```
VERSION     CQ268963.1  GI:41241567
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 17224 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 4
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||| ||| ||||||||  ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||| |||||||||||||| |||||||| ||||||| ||  ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792


RESULT 217
CQ305991
LOCUS       CQ305991                792 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 17096 from Patent WO0186003.
ACCESSION   CQ305991
VERSION     CQ305991.1  GI:41266568
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 17096 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN LUNG, SIGNAL = 3.9
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| ||||||||||||||||| ||| ||||||||  ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             ||||||||| |||||||||||| |||| |:|||||  |||||||||||| ||||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417
```

```
Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||||| |||||||||||||||| ||||||||| ||||||| ||   ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792
```

```
RESULT 218
CQ343148
LOCUS       CQ343148                792 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 17242 from Patent WO0157275.
ACCESSION   CQ343148
VERSION     CQ343148.1  GI:41292219
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 17242 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC002346.1
                     EXPRESSED IN BRAIN, SIGNAL = 6.2
                     NT HIT: AF208161.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU138405.1, EVALUE 0.00e+00
                     SWISSPROT HIT: P10269, EVALUE 2.00e-16"

  Query Match             9.5%;  Score 712.9;  DB 1;  Length 792;
  Best Local Similarity   93.8%;  Pred. No. 0;
  Matches  746;  Conservative    5;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       5955 AGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCCC 6014
             |||||||||  |||| |||||||||||||||| ||| |||||||| ||||| || ||||
Db          1 AGAAAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCC 60

Qy       6015 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGT 6074
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db         61 CTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTGGCCTGGT 120

Qy       6075 AAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 6134
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCCTATTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAA 180

Qy       6135 CTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACA 6194
             |||||||| |||||||||||| |||| |:|||||  |||||||||||  |||||||||
Db        181 CTGTTGGATGTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240

Qy       6195 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTC 6254
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||
Db        241 ATGGAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTC 297

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db        298 CAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGA 357

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        358 CACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACC 417

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        418 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGA 477

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||| |||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||
Db        478 ATCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 537

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        538 ATACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 597

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             |||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        598 TGGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCA 657

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||||| |||||||||||||||| ||||||||| ||||||| ||   ||||
Db        658 GTTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCT 717

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        718 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 777

Qy       6735 TTTAGACTYGCTAAC 6749
             ||||||||:||||||
Db        778 TTTAGACTTGCTAAC 792
```

```
RESULT 219
BV208491
LOCUS       BV208491                795 bp    DNA     linear   STS 10-DEC-2004
DEFINITION  ERVWE1_2140 Rhesus macaque genomic DNA Macaca mulatta STS genomic
            clone MMA2140, sequence tagged site.
ACCESSION   BV208491
VERSION     BV208491.1  GI:49533174
KEYWORDS    STS.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   Spindel,E.R., Pauley,M., Jia,Y., Thompson,S., Lankhorst,M.,
            Gravett,C., Lupo,S.L., Tchourbanov,A., Ali,H., Ojeda,S.R. and
            Norgren,R.B.
  TITLE     Targeted amplification of the 3' end of rhesus macaque orthologs of
            human genes
  JOURNAL   Unpublished (2004)
COMMENT
            Contact: Spindel ER
            Division of Neuroscience
            Oregon National Primate Research Center
            505 NW 185th Avenue, Beaverton, OR 97006, USA
            Tel: 403-690-5388
            Fax: 503-690-5384
```

```
          Email: spindele@ohsu.edu
          Primer A: aagagctttagacttgctaac
          Primer B: tagagtgaaatagcatgaaaa
          STS size: 795
          PCR Profile:
               Hot Start:       95 degrees C for 2.00 min
               Denaturation:    95 degrees C for 0.50 min
               Annealing:       51 degrees C for 0.50 min
               Polymerization:  72 degrees C for 1.00 min
               PCR Cycles:      35
               Extension        72 degrees C for 7.0 min
               Thermal Cycler:  MJ Instruments PTC100
          Protocol:
                    Template:        200 ng
                    Primer:          each 1uM
                    dNTP's:          each 200 uM
                    Taq Polymerase:  0.05 units/ul (Fast Start High
          Fidelity, Roche)
                    Total Vol:       50 ul

          Buffer:
                   MgCl2:           1.8 mM
                   Fast Start polymerase reaction buffer (Roche)

          Bases 5-795 are 94% homologous (Blast) to bases 2103-2892 of
          NM_014590.2.   Primers were chosen to amplify genomic DNA in the 3'
          region of ERVWE1.  As human sequence was used to design the
          primers, the primer sequences are not included in the rhesus
          sequence provided below.   To obtain additional information
          regarding primers or clones contact:  Dr. Robert Norgren; Dept of
          Genetics, Cell Biology & Anatomy; University of Nebraska Medical
          Center; 986395 Nebraska Medical Center; Omaha, NE  68198.  Email:
          rnorgren@unmc.edu
          A database containing sequences associated with this project can be
          found at:  http://rhesusgenechip.unomaha.edu/index.html.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Macaca mulatta"
                     /mol_type="genomic DNA"
                     /strain="Indian orgin"
                     /db_xref="taxon:9544"
                     /clone="MMA2140"
                     /clone_lib="Rhesus macaque genomic DNA"
                     /dev_stage="Adult"
                     /note="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid;
                     STS was amplified from rhesus genomic DNA with the human
                     forward and reverse primers listed above and subcloned
                     into pGEM-T Easy"
     gene            1. .795
                     /gene="ERVWE1"
                     /note="endogenous retroviral family W, env(C7), member 1"
     STS             <1. .>795
                     /gene="ERVWE1"

  Query Match             9.3%;  Score 699.7;  DB 1;  Length 795;
  Best Local Similarity   91.9%;  Pred. No. 0;
  Matches  730;  Conservative   15;  Mismatches   48;  Indels    1;  Gaps    1;

Qy       6751 GCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 6810
             || ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCGAAAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 61

Qy       6811 TCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGAR 6870
             || |||||| ||| ||||||||||:||||||||  |||||||||||||||||:|||| |:
Db         62 TCTGGAATCATCAGTGAGAAAGTTAAAGAAATTAAAGATCGAATACAACGTAGAGCAAAG 121

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| |||||||||||||||||||||||||  |:|||||| ||| ||||||||||||
Db        122 GAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAATCAATGGATACCCCGGATTCTCCCC 181

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||:||
Db        182 TTCTTAGGACTTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAAC 241

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGGAGCCCAAG 7050
             |||||||||| |||||||||| |||||| ||||| |:|||||||||||||||||| ||
Db        242 CTCCTTGTTAAGTTTGTCTCTTCTAGAATCAAAGCTATAAAACTACAAATGGAGCCCCAG 301

Qy       7051 ATGCAGTCCAAGAC-TAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATC 7109
             ||||||||| ||| |||| ||||||||||||||||||||||||:|||||| ||||
Db        302 ATGCAGTCCATGACTTAAGTTCTACCGCAGACCCCTGGACCGGCCTGCTAGCCCATGATC 361

Qy       7110 TGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACT 7169
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||
Db        362 CGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAACTGCACAACCCCTACT 421

Qy       7170 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACT 7229
             |||||||||||||||||||||||||||||| ||: ||| |||||||||||| ||||||
Db        422 ACGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCGACCAACCTCCCCAATAGCACT 481

Qy       7230 TAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGA 7289
             |||||||||||||||| |||| ||| |||||||||||||||||||||||||    |
Db        482 TAGGTTTTCCTGTTGAGAGGGGGCACTAAGAGACAGGACTAGCTGGATTTCCTAGGACAA 541

Qy       7290 YTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTG 7349
             :|||||||||:|||| ||||:|||||||||||||||||||||||||||||||| ||||||||
Db        542 CTAAGAATCCCTAAGGCTAGCTGGGAAGGTGACCACATCCACCTTTAAACAGGGGGCTTG 601

Qy       7350 CAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACA 7409
             ||||||||:||||||| |||||||||||||||||||||||||||||| |||||||||| |||
Db        602 CAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTGATTAGGCAAAAACA 661

Qy       7410 GGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATC 7469
             | ||||||||||||||||||||:||||||:|||||||||||| |||||||||| ||||
Db        662 GAAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAAGGATC 721

Qy       7470 GGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTT 7529
             |||||||||||||| |  ||| ||| |||||| ||||||||||||||||||||||||||||
Db        722 GGGATATAAACCCAGGCATTCCAGCTGGCAACAGCAACCCCCTTTGGGTCCCCTCCCTTT 781

Qy       7530 GTATGGGAGCTCTG 7543
             ||||||||||||||
Db        782 GTATGGGAGCTCTG 795


RESULT 220
AB063618
LOCUS       AB063618                772 bp    mRNA    linear   PRI 26-DEC-2001
DEFINITION  Homo sapiens HERV-W pol element mRNA, isolate:HWP-FB10.
ACCESSION   AB063618
VERSION     AB063618.1  GI:18143408
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Isolation and Phylogeny of the Human Endogenous Retrovirus HERV-W
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 772)
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HWP-FB10"
                     /db_xref="taxon:9606"
                     /tissue_type="Fetal Brain"
     repeat_region   1. .772
```

/rpt_family="HERV-W pol element"

Query Match 9.2%; Score 689.4; DB 1; Length 772;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 701; Conservative 67; Mismatches 4; Indels 20; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACCGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        461 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        521 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
Db        581 CTTGTGGCTGTCAGACAACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAGT 640

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
Db        641 GCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTTTTCCAGACAATGAAGA 700

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
Db        701 AAAGATAAAACATAACTGTCAACAAGTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 760

Qy       3870 TYTAGARGTTCC 3881
Db        761 TTTAGAGGTTCC 772
```

RESULT 221
CQ073091
LOCUS CQ073091 944 bp DNA linear PAT 20-JAN-2004
DEFINITION Sequence 8891 from Patent WO0157278.
ACCESSION CQ073091
VERSION CQ073091.1 GI:41042960
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelialcells
JOURNAL Patent: WO 0157278-A 8891 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .944
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC007567.1
EXPRESSED IN HELA, SIGNAL = 1.7"

Query Match 9.0%; Score 674.5; DB 1; Length 944;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 847; Conservative 4; Mismatches 90; Indels 23; Gaps 11;

```
Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761
```

```
Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             ||||||||||||| | |||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||||| |||||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||||| |||||| |||||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
             | ||
Db        941 CGAA 944


RESULT 222
CQ103917
LOCUS       CQ103917                944 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 12776 from Patent WO0157272.
ACCESSION   CQ103917
VERSION     CQ103917.1  GI:41072968
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 12776 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .944
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 7.9"

  Query Match             9.0%;  Score 674.5;  DB 1;  Length 944;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  847;  Conservative    4;  Mismatches   90;  Indels   23;  Gaps   11;

Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || ||||||||||||||| |||| ||||||| ||
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||| | |   ||| ||||||||||||   |
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||| |||| ||||||||||| |
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:||||||||||||||||| || ||
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||| |||||||||||||| |||||
```

```
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
                |||||||||||||||||||||||||||| |||||| |||||||||| || || | |
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
             | |   |||||||  |||||||||||||||||||||||||||| ||  | ||||| || ||
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
              | |        |||   | | |||||||||| |||||||||| ||||||||||||
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             |||||||||||||| | |||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||||| |||||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||||| |||||| |||||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
             | ||
Db        941 CGAA 944


RESULT 223
CQ225945
LOCUS       CQ225945                944 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 12784 from Patent WO0157273.
ACCESSION   CQ225945
VERSION     CQ225945.1  GI:41209005
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 12784 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .944
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

/note="MAP TO AC007567.1
EXPRESSED IN ADULT LIVER, SIGNAL = 1.6"

Query Match 9.0%; Score 674.5; DB 1; Length 944;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 847; Conservative 4; Mismatches 90; Indels 23; Gaps 11;

```
Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || ||||||||||||||| |||| ||||||| ||
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             ||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||| | |   ||| ||||||||||||| |
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||| |||| |||||||||||| |
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:||||||||||||||| || ||
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||||| |||||||||||||| |||||
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592

Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
               |||||||||||||||||||||||||||||| |||||| |||||||||| || || | |
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
             | |   |||||||  |||||||||||||||||||||||||||||||  ||   | ||||| || ||
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
             | |      |||   |  | |||||||||| |||||||||| ||||||||||||
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
             |||||||||||||| | |||||||||||| | |||||||||||||||||| ||||||||
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
             || ||::|||||||||||||| |||||||||||||||| | ||||||||||:||||||
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
             |||| |||||||||||||||||||||||||| |||||| |||||||||||||| |||
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
             | ||
```

Db 941 CGAA 944

RESULT 224
CQ301087
LOCUS CQ301087 944 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 12192 from Patent WO0186003.
ACCESSION CQ301087
VERSION CQ301087.1 GI:41261664
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 12192 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .944
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC007567.1
EXPRESSED IN LUNG, SIGNAL = 1.3"

Query Match 9.0%; Score 674.5; DB 1; Length 944;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 847; Conservative 4; Mismatches 90; Indels 23; Gaps 11;

```
Qy       1906 TTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTG 1965
             |||||||| || || |||||| |||| || ||||||||||||||| |||| ||||||| ||
Db          1 TTGTCCAAATACAAATAAGCCACCCCCTCATCCATGCCCCTTATGTCAAAGGAATCAATG 60

Qy       1966 GAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGAT 2025
             ||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||
Db         61 GAAGGCCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGAT 120

Qy       2026 CCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGA 2085
             || |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||
Db        121 CCGGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAGCCCAACCCATCACCCTCACAGA 180

Qy       2086 GCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGC 2142
             |||| |||||||||||||||||||||||||| | |   ||| ||||||||||||| |
Db        181 GCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTTTCCTGGACACTGGCTC 240

Qy       2143 GGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCTGTCACTATTCT 2202
             || ||||| |||||||||| |||||||||||||||||| |||| |||||||||||| |
Db        241 GGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCTTCCACATCTGTCACTATCC- 299

Qy       2203 GAGGGGGTCCNTAAGACGGGCAGTCACTAGATACTTTYTCCCAGCCACTAAGTTATGAAC 2262
               | |||  | || ||| | |||||||| ||||| ||:||||||||||||||| || ||
Db        300 --GAGGGGTCCTAGGACAGCCAGTCACTTGATAC-TTCTCCCAGCCACTAAGTTGTG-AC 355

Qy       2263 TGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTG 2322
             |||||| ||||| || |||||||||||||||||||||||| |||||||||||||| |||||
Db        356 TGGGGAACTTTACTC-TTTCACATGCTTTTCTAATTATGCCTGAAAGCCCCACTCCCTTG 414

Qy       2323 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGA 474

Qy       2383 ACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 2442
             ||| |||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db        475 ACAACCGTTTGTGGT-CCCTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGA 533

Qy       2443 AGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCA 2502
             |||||||||||| ||| |||| ||||||| |||||||||| ||||||||||||||||
Db        534 AGGACAATATGGATGAG-CAAAAAATGCCCATCCTGTTCAAATTAAACTAAAGGATTCTG 592
```

```
Qy       2503 CTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAA 2562
Db        593 TCCCCTTTCCCTACCAAAGGCAGTACCCCCTTAGACCCGAGGCCCAACAAAGACTCAATA 652

Qy       2563 AGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTA 2622
Db        653 AAA---TTAAGGAACTAAAAGCCCAAGGCTTAGTAAAACCATTCAATAGTCCCTCCAATA 709

Qy       2623 ATTCCGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGA 2682
Db        710 CTCC-------AATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGA 761

Qy       2683 TCTCAGGATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATA 2742
Db        762 TCTCAGGATTATCAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATA 820

Qy       2743 CTGTGMYTTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATG 2802
Db        821 CTCTGCTTTCCCAAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATG 880

Qy       2803 CCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTT 2862
Db        881 CCTTTTTCTGCATCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTT 940

Qy       2863 CAAA 2866
Db        941 CGAA 944
```

RESULT 225
AB063621
LOCUS AB063621 787 bp mRNA linear PRI 26-DEC-2001
DEFINITION Homo sapiens HERV-W pol element mRNA, isolate:HWP-PL4.
ACCESSION AB063621
VERSION AB063621.1 GI:18143411
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kim,H., Jeon,S. and Lee,Y.
TITLE Isolation and Phylogeny of the Human Endogenous Retrovirus HERV-W
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 787)
AUTHORS Kim,H., Jeon,S. and Lee,Y.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2001) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .787
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="HWP-PL4"
/db_xref="taxon:9606"
/tissue_type="Placenta"
repeat_region 1. .787
/rpt_family="HERV-W pol element"

Query Match 8.8%; Score 660.7; DB 1; Length 787;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 687; Conservative 63; Mismatches 37; Indels 5; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGGTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGGACCAGGGCCCTCAGTGAGGAACATATCCAGTCTTTACTGGCTTATCCTTATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAGAGA-TTCCTTGGCATAACAGGTTTCTGCCGAGTATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        180 -CCCAGGTACGGTGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        239 CAATACCCATTTATTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGAGTTTTCTTTATAT 356

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        357 GTCACAG-AAAAAACAGGAATAGCTCTAGAAGTCCTTACACAGGTCTGAGGGATGAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        416 GCAACCCACGGCATACCTGAGTAAGGCAATCGATGTAATGGCAAAGGGTTGGCCTCATTG 475

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        476 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 535

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        536 AAGAGATCTTACTGTGTGGACTTCTCATGATGTGAACTGCATACTCACTGCTAAAGGAGA 595

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
Db        596 CTTGTGGCTGTCAGACAACCATTTATTTAAATATCAGGCTCTATTACTTGAAGGGCTAGT 655

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
Db        656 GCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGA 715

Qy       3810 AAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCT 3869
Db        716 ACAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACGCCACTCGAGGGGACCT 775

Qy       3870 TYTAGARGTTCC 3881
Db        776 TCTAGAGGTTCC 787
```

RESULT 226
BV469300/c
LOCUS BV469300 841 bp DNA linear STS 06-APR-2005
DEFINITION G591P609217FF10.T0 Clint Pan troglodytes verus STS genomic, sequence tagged site.
ACCESSION BV469300
VERSION BV469300.1 GI:62228728
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 841)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA

Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 841
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .841
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>841

Query Match 8.8%; Score 659.4; DB 1; Length 841;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 725; Conservative 50; Mismatches 66; Indels 12; Gaps 5;

```
Qy       2520 AGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTA 2579
             |||||||||||||| |||||  |||||||||||||| |||||||||||||||||||| ||
Db        841 AGGCAGTACCCCCTTAGACCTGAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTA 782

Qy       2580 AAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGA 2639
             ||| | |||||| |||||||| ||||||   |  |||||    || | |          ||
Db        781 AAATCTCAAGGCCTAGTAAAAGCATGCAATAGCCCCTGTGATACTCC-------ATTTTT 729

Qy       2640 GGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATG 2699
               | | ||||||||||    ||||||||| ||||||||||||||||||||||||||||||
Db        728 AGGAGTACAGAAACCCAACAGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT- 671

Qy       2700 GAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATA 2759
             |||||| |||||| | |||||||||||||||||| |||||||||| ||::||||||||||
Db        670 GAGGCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCCAAATA 611

Qy       2760 CCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCT 2819
             |||||||||||||||| |||||:||||||||||:|||||||||  ||||||||||
Db        610 CCAGAGGAAGCAGAGTGTTTTACAGTCCTGGACCTTAAGGATGCCTTTGTCTGCATCCCT 551

Qy       2820 GTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAA 2879
             |||||||||| |||||||||||||||||||||||||| |||  ||| ||:  |||||
Db        550 GTACATCCTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACTCAATGTCTCAC 491

Qy       2880 CTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCA 2939
             ||||||||||:|||||||||||||||||||||:|| |||||||||||||||||
Db        490 CTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCACATCTATTTGGCCAGGCA 431

Qy       2940 TTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGAT 2999
             ||||||||||||||||:||:|::|||||||||||||||||||||:|||:||||||
Db        430 TTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACACTCTTGTCCTTCAGTAGGTGGAT 371

Qy       3000 GATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTM 3059
             |||||||||||:||: ||::|||||||||||||||||||||||||||||||| :|||||:
Db        370 GATTTACTTTTAGCCACCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGTGCTCTTA 311

Qy       3060 AATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAG 3119
             ||||||||||:|||||||||||||:||||||||||:|:|:|||||:|||| |||||||
Db        310 AATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTTCTCACAG 251

Qy       3120 CAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRY 3179
             |||||||||||||||||:||| : |||||| |||||||||:||||||||||||||||:::
Db        250 CAGGTTAAATACTTAGGGCTACAGTTATCCCAAGGCACCAGGGCCCTCAGTGAGGAATGT 191

Qy       3180 ATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT 3239
             |||||||||||||||||||||||||||||:|||||||||||||||||||||:|::||||||
Db        190 ATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGAGGGTTCCTT 131

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
             |||:|||:|||:||||||| ||:||||||| |||| ||:  ||:|||||||||| :||||
Db        130 GGCATAACAGGTTTCTGCCAAATATGGATT-CCCAAGTACAGCAAAATAGCCAGACCATT 72

Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGA 3359
             |:|||||: |||||||||||||||  |||  | |   |||||||:|||||||||:|:||||
Db         71 ATATACACCAATTAAGGAAACTC--AAAAGCCAATACCATTTAGTAAGATGGACACCTGA 14

Qy       3360 AGYMRAAGTGGCT 3372
             ||:::|||  |||
Db         13 AGCAGAAGCAGCT 1
```

RESULT 227
BV457035
LOCUS BV457035 824 bp DNA linear STS 06-APR-2005
DEFINITION qsr62c05.b1 Clint Pan troglodytes verus STS genomic, sequence tagged site.
ACCESSION BV457035
VERSION BV457035.1 GI:62216463
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 824)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 824
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI

```
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .824
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>824

  Query Match             8.7%;  Score 655;  DB 1;  Length 824;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  703;  Conservative   39;  Mismatches   67;  Indels    6;  Gaps    3;

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
              | | | ||:| |||||||||||| |||||||||| |||:||:||||||||||||||:|
Db         13 TTATCTAGCCACATTTCTTCCAGAAAATGAAGAAAAAATAGAACATAACTGTCAACAAGT 72

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
              |||| |||||||||| |||||||||||||||||:||||:|||||:|||||||| |:||
Db         73 AATTGCTCAAACCTACGCCACTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATTCCGA 132

Qy       3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
              |||||||||||| |||||||||||||  ||| |||||||||||||  |||||:|    |
Db        133 -CTTCAACTTGTACACTGATGGAAGTTTATTTTTAGAAAAAGGACTTTAAAAAGCGTATT 191

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
              |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db        192 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 251

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
              ||:||||:|||||| ||||||||:|||:|  ||  |||||||||||||| : |||||||
Db        252 CTCAGCTGGCAGAAATAATAGCCCTCACTCAGGTGCTAGAATTAGGAGAAAGGAAAAGGG 311

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
              :|||||||||||||||||:|:||||||:|||||||| ||||||||||||::|||||||
Db        312 TAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 371

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
              |:|:|||||||||||||||||||:|| :||  ||||||||:|||||||||||||||||
Db        372 GGAGAGAAAGGGAATTCCTAACTTCTGAAGGACAACCTATCAAACATCAGGAAGCCATTA 431

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
              |||:|||||||:||||:| ||||||| |||:||||||||:| ||||||||| |:||||||
Db        432 GGAGATTATTATTGGCTGCACAGAAATCTAAAGAGGTGGCAATCTTACACTACTGGGGTC 491

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
              |||| |||||||||:||||||||||:|||:|||: ||||||  :||||||||||:||||
Db        492 ATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGTGGATATTGAAGCCAAAA 551

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
              ||||  |||||||||||||||||| ||||||||||||  ||   ||| |||||| |||||
Db        552 GAGCCACAAGGCAGGACCCTCCATAAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTA 611

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
              ||||| ||| |||||||||||||||||||||||| |||| | ||||||||||||||||||
Db        612 ATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCAAGAGAGAGAGAATGGGGAACCTCAC 671

Qy       4437 GAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGC 4493
              |||    ||||| |||||||| ||| || |||||||||||||||||||||||||||||||
Db        672 AAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACTGAAGAAGGGAAAATACTTTTGT 731

Qy       4494 CTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCG 4553
              ||||| ||| || ||||||||   || |||||||||| |||||||||||||||||||| |
Db        732 CTGCAGCTAACCCATGGAAAT--ATTTAAACCCTTCACCAAACCTTTCACTTAGGCATTG 789

Qy       4554 ATAGCACCCATCARATGGCCAAATCATTATTTACT 4588
              |||||||||||||:|||||||||| ||||||||||
Db        790 ATAGCACCCATCAGATGGCCAAATTATTATTTACT 824


RESULT 228
AY836220/c
LOCUS       AY836220                749 bp    DNA     linear   PRI 20-SEP-2005
DEFINITION  Homo sapiens endogenous virus HERV-W clone W6 pol gene, partial
            sequence.
ACCESSION   AY836220
VERSION     AY836220.1  GI:56684744
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Forsman,A., Yun,Z., Hu,L., Uzhameckis,D., Jern,P. and Blomberg,J.
  TITLE     Development of broadly targeted human endogenous gammaretroviral
            pol-based real time PCRs: Quantitation of RNA expression in human
            tissues
  JOURNAL   J. Virol. Methods 129 (1), 16-30 (2005)
   PUBMED   15967513
REFERENCE   2  (bases 1 to 749)
  AUTHORS   Forsman,A., Yun,Z., Uzhameckis,D., Hu,L., Jern,P. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2004) Section of Virology, Dept of Medical
            Sciences, Uppsala University, Academic Hospital, Uppsala, Uppland
            751 85, Sweden
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="W6"
                     /note="endogenous_virus: HERV-W"
     gene            complement(<1. .>749)
                     /gene="pol"
                     /note="5' end"

  Query Match             8.7%;  Score 653;  DB 1;  Length 749;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  710;  Conservative    7;  Mismatches   22;  Indels    8;  Gaps    5;

Qy       4737 CAAGCCCAAACCTCAGGGATTT--CAGTATCTACTAGTCT-GGGTARATACTTTCACGGG 4793
              || ||| || | ||||| ||||    || | ||||||| | ||||| | ||||||  |||
Db        747 CANGCCNAACCNTCAGGAATTTTCAGGTTTTTACTAGTTTGGGGTANANACTTTCNNGGG 688
```

```
Qy       4794 TTGGG---CARAGGCCTTCCCCTGTAGGACAGAAAAGGCCC-AAGAGGTAATAAAGGCAC 4849
Db        687 TTGGGGCAAAGGGCNTTCCCCCTNTAGGNCAGAAAAGGCCCNAAGAGGTAATAAAGGCAC 628

Qy       4850 TAGTTCATG-AAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGC 4908
Db        627 TAGTTCATGAAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGC 568

Qy       4909 CCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCAC 4968
Db        567 CCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCAC 508

Qy       4969 TTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTC 5028
Db        507 TTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTC 448

Qy       5029 AAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGCCT 5088
Db        447 AAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCT 388

Qy       5089 ATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATG 5148
Db        387 ATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATG 328

Qy       5149 CTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAG 5208
Db        327 CTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTTAG 268

Qy       5209 TTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCT 5268
Db        267 TTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAACCT 208

Qy       5269 ATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTCCC 5328
Db        207 ATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCC 148

Qy       5329 TTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 5388
Db        147 TTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTAT 88

Qy       5389 YTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 5448
Db         87 CTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCA 28

Qy       5449 AATCCTGGATACTGCCAAAGGAACCTG 5475
Db         27 AATCCTGGATACTGCCAAAGGAACCTG 1
```

```
RESULT 229
BV598358
LOCUS       BV598358                808 bp    DNA     linear   STS 14-APR-2005
DEFINITION  S217P6430RD2.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV598358
VERSION     BV598358.1  GI:62537272
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 808
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>808

  Query Match             8.6%;  Score 643.5;  DB 1;  Length 808;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  703;  Conservative   68;  Mismatches   33;  Indels   21;  Gaps    7;

Qy       2791 ACCTTMAGGATGCCT-TCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCC 2849
Db          4 ACCTTAAGGATGCCTCTTTTCTGCATCCCTGTAC-------------TTCTTGTTTGCC 49

Qy       2850 TTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTC 2909
Db         50 TTTGAAGATCCTTCAAACCCAACATCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTC 109

Qy       2910 AGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATAC 2969
Db        110 AGGGATAGCCCCCATCTATTTGGCCAGGAATTAGCCCAAGACTTGAGCCAGTTCTCATAC 169

Qy       2970 CTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACC 3029
Db        170 CTGGACACTCTTGTCCTTCGGTATGTGGATGATTTACTTCTAGCCACCCGTTCAGAAACC 229

Qy       3030 TTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTT 3089
Db        230 TTGTGCCATCAAGCCACCCAAGCGCTCTTTAATTTCCTCGCCACCTGTGGCTACAAGGTT 289

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
```

```
              |||||||:|:|:|||||:|||||||||||| |||||||||||||| ||:||||:|||||||
Db        290 TCCAAACCAAAGGCTCAGCTCTGCTCACAACAGGTTAAATACTTCGGGCTAAAATTATCC 349

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||| |||||:|||||||||||||| :::|||||||| |||||||||||||||||||:|
Db        350 AAAGACACCAGGGCCCTCAGTGAGGACTGTATCCAGCCCATACTGGCTTATCCTCATCCC 409

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:| :|||||||||:|||:|||:||||||||||:|||||||
Db        410 AAAACCCTAAAGCAACTAAGAGCATTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 469

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||||:||||| |||||:|||||||||||||||||||||
Db        470 -CCCAGGTATGGCAAAATAGCCAGGCCATTACATACACTAATTAAGGAAACTCAGAAAGC 528

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:||||||| |:|: |||||:::||| ||||||| ||||||  ||  |
Db        529 CAATACCCATTTAGTAAGATGAACACTTGAAGCAGAAGCGGCTTTCTAGGCCCTAAAGAA 588

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  |||||||:||||||||||||:|||| || :|||||||||||||:||:|||:
Db        589 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCTAATGGGGCAAGACTTTTCTTTATAT 646

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||  ||||||||:||:||||||:|||||||||||||| |:|||:|||||:||||||
Db        647 GTCACAG-GAAAAACAGAAATAGCTCTAGGAGTCCTTACAC-GGTCCGAGGGATGAGCTT 704

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              || |||:: |||::||||||:||||||||:|||||||||||||||||:|:    ||
Db        705 GCTACCCAGGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCCTCATG 764

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGT 3614
              ||||: |||||||| ||||||||||||::|||||||||||||||
Db        765 TTTAT-GGTAGTGATGGCAGTAGCAGTCTTAGTATCTGAAGCAGT 808
```

```
RESULT 230
AY836221/c
LOCUS       AY836221                751 bp    DNA     linear   PRI 20-SEP-2005
DEFINITION  Homo sapiens endogenous virus HERV-W clone W4 pol gene, partial
            sequence.
ACCESSION   AY836221
VERSION     AY836221.1  GI:56684745
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Forsman,A., Yun,Z., Hu,L., Uzhameckis,D., Jern,P. and Blomberg,J.
  TITLE     Development of broadly targeted human endogenous gammaretroviral
            pol-based real time PCRs: Quantitation of RNA expression in human
            tissues
  JOURNAL   J. Virol. Methods 129 (1), 16-30 (2005)
   PUBMED   15967513
REFERENCE   2  (bases 1 to 751)
  AUTHORS   Forsman,A., Yun,Z., Uzhameckis,D., Hu,L., Jern,P. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2004) Section of Virology, Dept of Medical
            Sciences, Uppsala University, Academic Hospital, Uppsala, Uppland
            751 85, Sweden
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="W4"
                     /note="endogenous virus: HERV-W"
     gene            complement(<1. .>751)
                     /gene="pol"
                     /note="5' end"
```

```
  Query Match             8.3%;  Score 624.6;  DB 1;  Length 751;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  697;  Conservative    8;  Mismatches   36;  Indels    8;  Gaps    6;

Qy       4735 CACAAGCCCAAACCTCAGGGATTTCA-GTATCTACTAGTCTGGGTARATACTTTCACGGG 4793
              | |||| |||| ||| ||||||| || || |  ||||||  |||||:|| ||| || |||
Db        749 CCCAAGNCCAANCCTNAGGGATTNCAGGTTTTNACTAGTTGGGGTAGATNCTTNCAGGGG 690

Qy       4794 TTGGGCARAGGCCTTCCCCTGTAGGACAGAAAA---GGCCCAAGAGGTAAT-AAAGGCAC 4849
              | ||   |: | ||||||||| |  | |  | ||   | || | |||||||| ||||||||
Db        689 TGGGCAAAGGCCCTTCCCCTTTAGGCCAANAAAGGGCCCAANGAGGTAATAAAAGGCAC 630

Qy       4850 TAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAG-AGTGACAATAGC 4908
              | |||||||||||||||| |||||||||||||| |||| |||||  || |||||||||
Db        629 TGGTTCATGAAATAATTCCNAGATTCGGACTTCCCNGAGGNTTACANAAGGGACAATAGC 570

Qy       4909 CCTGCTTTCCAGGCCACAGTAACCC-AGGGAGTATCCCAGGCGTTA-GGTATACGATATC 4966
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||
Db        569 CCTGCTTTCCAGGCCACAGTAACCCAAGGGAGTATCCCAGGCGTTAGGGTATACGATATC 510

Qy       4967 ACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYAC 5026
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||:||
Db        509 ACTTACACTGCGCNTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACAC 450

Qy       5027 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTGC 5086
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        449 TCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGC 390

Qy       5087 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 5146
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 CTATAGCCTTAAAAAGAATCTGCAACTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAA 330

Qy       5147 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 5206
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAGACAGCCAACTT 270

Qy       5207 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 5266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGGAAC 210

Qy       5267 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTAGTCAAGTC 5326
              ||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db        209 CTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTC 150

Qy       5327 CCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 5386
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CCTTCCCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTT 90

Qy       5387 ATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 5446
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGT 30

Qy       5447 CAAATCCTGGATACTGCCAAAGGAACCTG 5475
              |||||||||||||||| ||||||||||||
Db         29 CAAATCCTGGATACTGCGAAAGGAACCTG 1
```

```
RESULT 231
BV552074
LOCUS       BV552074                775 bp    DNA     linear   STS 09-APR-2005
DEFINITION  S221P61492RD1.T0 Yvonne Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV552074
VERSION     BV552074.1  GI:62443094
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
```

```
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 775
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Yvonne"
     STS             <1. .>775

  Query Match             8.3%;  Score 621.5;  DB 1;  Length 775;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches  641;  Conservative   53;  Mismatches   71;  Indels    1;  Gaps    1;

Qy       3513 ACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTT 3572
             |||::||||::| || |:|| |||||: |||||||||| |||||||:|:||||||||
Db         11 ACCTGTGGCATAGCTAAGTAGGGAAATAGATGTAGTGGCTAAGGGTTGGCCTCATTGTTT 70

Qy       3573 AYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAG 3632
             |:|||||||||  ||||||||||::|||| |||||||||||||||||||||||||||:||
Db         71 ACGGGTAGTGGCAGCAGTAGCAGTCTTAGTGTCTGAAGCAGTTAAAATAATACAGGGAAG 130

Qy       3633 AGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTT 3692
              |||||||||| ||| |||||| ||||:|||||::||||||||||||||||||||||||
Db        131 GGATCTTACTGCGTGAACATCTTATGATGTGAATGGCATACTCACTGCTAAAGGAGACTT 190

Qy       3693 GTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCT 3752
             |||||||| |||||||:|||||||||||:|||||||||||||||| ||:||||||||||
Db        191 GTGGCTGTGAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTAAAGGGCCAGTGCT 250

Qy       3753 GCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAA 3812
             ||:|||| |||||||||||||||||||||:| |||||||| |||||||||||||||||
Db        251 GCAACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTTCAGACAATGAAGAAAA 310

Qy       3813 GATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYT 3872
             ||||:||:|||||||||||||:||||| |||||||||||  | || ||||||||||:|
Db        311 GATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGGCGCTAGAGGGGACCTTTT 370

Qy       3873 AGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAG 3932
             |||: ||||:||||||||||:|||| |||||| |||||||||| ||||| | |||
Db        371 AGAGTTTCCCTTGACTGATCCTGACC-TCAACTCGTATACTGATTGAAGTCCTTTTTCTA 429

Qy       3933 AAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTA 3992
              |||   ||||| |||||:| |||||||||||||||||||||||||||:|| ||||||
Db        430 CAAAGAAACTTCTAAAAGTGAGGTATGCAGTGGTCAGTGATAATGGAATACTTAAAAGTA 489

Qy       3993 ATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCAC 4052
             |||||||||||||| ||||||||||:||||:|||||||||||||||:|||:|:||||||
Db        490 ATCCCCTCACTCCAAGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTTGGGCAC 549

Qy       4053 TAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAG 4112
             |||||||||||||| :||||| ||:|||||||||||||| ||||:|:||||||:|||||
Db        550 TAGAATTAGGAGAAAGAAAAAAGGTAAATATATATACAGCCTCTAAGTATGCTTACCTAC 609

Qy       4113 TCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACAC 4172
             || ||||||||||||:: |||| | :|:||||| |||||||||||||||:   :|||||
Db        610 TCCTCCATGCCCATGCAACAATGTAGAGAGAAAAGGAATTCCTAACTTCCCGAGGAACAT 669

Qy       4173 CTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGG 4232
             |||||:|||||| |||||||||||||:|| ||||:||||:|||||||||||  :| |||
Db        670 GTATCAAACATCAAGAAGCCATTAGGAGATAATTATTGGCTGTACAGAAACCCTAAAAGG 729

Qy       4233 TGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGA 4278
             |  :|    ||||||||:  ||||| |   |||||||:|| || |
Db        730 TAAAACATCTACACTGCCAAGGTCACCCAGAAGGAAAGAAAGGGAA 775


RESULT 232
AB063620
LOCUS       AB063620                783 bp    mRNA    linear   PRI 26-DEC-2001
DEFINITION  Homo sapiens HERV-W pol element mRNA, isolate:HWP-PL3.
ACCESSION   AB063620
VERSION     AB063620.1  GI:18143410
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Isolation and Phylogeny of the Human Endogenous Retrovirus HERV-W
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 783)
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-2001) Heui-Soo Kim,.Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HWP-PL3"
                     /db_xref="taxon:9606"
                     /tissue_type="Placenta"
     repeat_region   1. .783
                     /rpt_family="HERV-W pol element"
```

```
  Query Match             8.3%;  Score 619.9;  DB 1;  Length 783;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  672;  Conservative   66;  Mismatches   43;  Indels   13;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||| ||||||:||| ||||| ||||||:::|||||||||||| ||||||||||||||:|
Db         61 AAAAGCACCAAGGCGCTCAGCGAGGAACGTATCCAGCCTATATTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||:|::|||| ||||:|||:|||:|||||| |||:|||||||
Db         121 AAAACCCTAAAGCAACTAAGAGGGTTCCGTGGCATAACAGGCTTCTGCTGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:||||:|||||||||||:|||||:|||||:||||||||||||||||||||||
Db         181 -CCCAGGTATGGCGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:| ||||| |: :|| |||:::||| |||||||||||||   || |
Db         240 CAATACCCATTTAGTCAGATGTACGCCTAAAGCAAAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:||||||||||||:|||||||||:|| ||||||||||:||:| |:
Db         300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGCCAAGACTTTTCTTTACAT 357

Qy       3450 RTCACAG-AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             :| |||| ||||||||||:||:||||||:|||||||||||||||:|||:|||||: ||||
Db         358 GTAACAGAAAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACCAGCT 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::||||||:||||||||:||||||||||||||||||| :|:|||||
Db         418 TGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTAGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||:||| |||||  |||||||||||::||| ||||||||||||||||||||||||||
Db         478 GTTTATGGGCAGTGGCAGCAGTAGCAGTCTTAGGATCTGAAGCAGTTAAAATAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |:|||||||||||||||||||||||||| |:|| ||::|||||||||||||||| |||
Db         538 GAAGAGATCTTACTGTGTGGACATCTCATAATGTAAACAGCATACTCACTGCTAAAAGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAG 3748
             |||||||||||||||||: ||| ||||||:||||||||||||    ||: |||||
Db         598 ACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTA----AAGAGCCAG 653

Qy       3749 TGCTGCRACTGTGCACTTGTGCAACTCTTAA-CCCAGYCNCATTTCTTCCAGACAATGAA 3807
             ||||||:|||||||||||||||||||||||| |||||:| |||||||||||||||| ||
Db         654 TGCTGCAACTGTGCACTTGTGCAACTCTTAATCCCAGCCACATTTCTTCCAGACAATAAA 713

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              |||||    ||:|||||||||||||| :| ||| |||| ||||| ||| ||  |||||||
Db         714 AAAAAG----AACATAACTGTCAACAGGTGATTGCTCACACCTACGCCGCTTAAGGGGAC 769

Qy       3868 CTTYTAGARGTTCC 3881
              |||:||||:|||||
Db         770 CTTCTAGAGGTTCC 783


RESULT 233
BV470413/c
LOCUS       BV470413                 759 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P612187FD2.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV470413
VERSION     BV470413.1  GI:62229841
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 759
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>759

  Query Match             8.2%;  Score 617.3;  DB 1;  Length 759;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  692;  Conservative    5;  Mismatches   57;  Indels    7;  Gaps    4;

Qy       5657 GCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAA 5716
             ||||||  |||||||||||| |||:|||:|||||||  || ||||||| || ||||||||
Db          759 GCTGTACAACCAGTAGCTCACCTTACCAAGAGTTTC-CTGAAGAATGCGGCTTCCCGGAA 701

Qy       5717 ATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACA 5776
             ||||||||| |||||||||||||||| | |:||| |||| ||||||||||  |||||||
Db          700 ATATTGATGTCCCATCGTATAGGAGTATATCTAAAGGAAACCCCACCTTCATGGCCCACA 641

Qy       5777 CCCATATGCCCCGCAA---CTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTC 5833
```

```
                |||| ||||||| |||     ||||| ||||||| |||||||||||||||||||||||||
Db        640 CCCACATGCCCCACAAGAGTGGCTATAACTCTGCAACTCTTTGCATGCATGCAAATACTC 581

Qy       5834 ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCT 5893
                |||||||||||| |||||||||||||||||||||||||||||||||||||| ||  ||||
Db        580 ATTATTGGACAGGGAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGCCA--GTCT 523

Qy       5894 GTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAA 5953
                |||||||||||||||||| |  |||||||||||||||||||||||||||||||||||||
Db        522 GTTGGACTTACTTCACCCATAACGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAA 463

Qy       5954 GAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCSGGGTACATGGCACCTCTAGCC 6013
                |||||||||| |||| ||| |||||||||||| ||| |||||||  | ||| | ||||
Db        462 GAGAAAAACACATAAAGGAAATAATCTCCCAACTGACCTGGGTACAAAGTACCCCCAGCC 403

Qy       6014 CCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCGTACCCATACTCGCCTGG 6073
                |||||||||||||||||||||||||||||||||||||||||| ||||||||||| ||||
Db        402 CCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCTCCATACCCATACTCGGCTGG 343

Qy       6074 TAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTA 6133
                ||||| |||||||||| ||||||||| |||||||| |||| ||||||||||||||||||
Db        342 TAAGCTTATTTAATACTACCCTCACTAGGCTCCATAAGGTTTCGGCCCAAAACCCTACTA 283

Qy       6134 ACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAAC 6193
                |||||||||| |||||||||||| ||||||:|||||   ||||||||||||||||||||
Db        282 ACTGTTGGATGTGCCTCCCCCTGCACTTCAGGCCATA-CATTCAATCCCTGTACCTGAAC 224

Qy       6194 AATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTT 6253
                |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||
Db        223 AATGGAACAACTCCAGCACAAAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTT 164

Qy       6254 CCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACAT 6313
                |||||:|||| |||||||||||||||||| |||||||||||||| |||||||| |
Db        163 CCAATCTGGAGATAACCCATACCTCAAACCTTACCTGTGTAAAATTTAACAATACTATAG 104

Qy       6314 ACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTAC 6373
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 ACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTAC 44

Qy       6374 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGT 6414
                |||||||||||| || |||||||||||||||| |||||||
Db         43 CCTCAGGAATATATTGTGTCTGTGGTACCTCAGACTATCGT 3
```

```
RESULT 234
BV627795
LOCUS       BV627795                765 bp    DNA     linear   STS 15-APR-2005
DEFINITION  S216P60347RA10.T0 Masuku Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV627795
VERSION     BV627795.1  GI:62617973
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 765
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Masuku"
     STS             <1. .>765
```

```
  Query Match             8.2%;  Score 613.5;  DB 1;  Length 765;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  650;  Conservative   62;  Mismatches   53;  Indels    5;  Gaps    4;

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
                |||||||:|||||||||||||||||: :  |   |||||||||||||||||| |||:|||
Db          1 AGGCACCAGGGCCCTCAGTGAGGAACCCCCCTTTCCTATACTGGCTTATCCTTATCCCAA 60

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
                |||||||||||    || |:|::|||||||||:|||:|||:||||||||||:||||||| |
Db         61 AACCCTAAAGC-CTTATGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT-C 118

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
                ||||||:||  :||| |||||| :|||||:|||||:|||||| ||||||||||||||||
Db         119 CCAGGTATGATGAAACAGCCAGACCATTATATACACTAATTAAAGAAACTCAGAAAGCCA 178

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
                ||||||||||:|||||||||:|:||||||:::||| |||||||||||||||  ||  | |
Db         179 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAGG 238

Qy       3392 GCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRT 3451
                ||||  |||||||:|||||||||||:|||||||:|||||||||||||:||:| |::|
Db         239 CCCTT--ACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTACACGT 296

Qy       3452 CACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGC 3511
                |||||||||||||||:||:||||||:|||||||||||||||:|||:| |||: |||||||
Db         297 CACAGAAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAAGGACAAGCTTGC 356

Qy       3512 AACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTT 3571
                ||||::||||::||||||:|||||||||:|| |||||||||||||| |||:|:|||||| |
```

```
Db        357 AACCCATGGCATACCTGAGTAAGGAAATTGTTGTAGTGGCAAAGGTTTGGCCTCATTGAT 416

Qy       3572 TAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRA 3631
              ||:|||||||||  ||||||||||::|||||||||||||||||||||||||||||||:|
Db        417 TATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAA 476

Qy       3632 GAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACT 3691
              |||||||||||||||| |||||||  ||||:|||||::||| |||||||||||||||||||
Db        477 GAGATCTTACTGTGTAGACATCTAATGATGTGAACGGCACACTCACTGCTAAAGGAGACT 536

Qy       3692 TGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGC 3751
              |||||||||||||||||: ||||||||||:|||||||||||||||||:| |||||||
Db        537 TGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCAGTGC 596

Qy       3752 TGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAA 3811
              || :|||||||||||||||| ||||||||||:| |||||||| ||||||||||||||||
Db        597 TGTGACTGTGCACTTGTGCACCTCTTAACCCAGCCACATTTCTTACAGACAATGAAGAAA 656

Qy       3812 AGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTY 3871
              |||||:||:|||||||||     |:||||| |||||||||||||| |||  |||||||||:
Db        657 AGATAGAACATAACTGTGTCACAATAATTGCTCAAACCTATGCCGCTCAGGGGGACCTTT 716

Qy       3872 TAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAG 3921
              ||||:|||||:|||||||||||: |||    || ||||||||||||||||
Db        717 TAGAGGTTCCCTTGACTGATCCCCACCCCTCAC-TGTATACTGATGGAAG 765


RESULT 235
BV500243
LOCUS       BV500243                711 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S222P62744RB12.T0 Karlien Pan troglodytes verus STS genomic,
            sequence tagged site.
ACCESSION   BV500243
VERSION     BV500243.1  GI:62343985
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 711
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Karlien"
     STS             <1. .>711

  Query Match             8.1%;  Score 608.1;  DB 1;  Length 711;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  618;  Conservative   45;  Mismatches   47;  Indels    1;  Gaps    1;

Qy       3579 AGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCT 3638
              |||||  |||||||||||:: |||||||||||||||||||||| |||  |||:||||||||
Db          2 AGTGGCAGCAGTAGCAGTCTCAGTATCTGAAGCAGTTAAAATGATATTGGGAAGAGATCT 61

Qy       3639 TACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCT 3698
              |||||||||||||||| |||:|||||::||||||||| |||||||| ||||||||||||
Db         62 TACTGTGTGGACATCTCGTGATGTGAATGGCATACTCATTGCTAAAGAAGACTTGTGGCT 121

Qy       3699 GTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACT 3758
              ||||||||||:|||| |||||||:|||||||||||||||:| ||||||||||:|||
Db        122 GTCAGACAACCGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGACCAGTGCTGCGACT 181

Qy       3759 GTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARA 3818
              |  ||| |||||||||||||| | ||:| |||||||||||| |||||||||| |||:|
Db        182 GCACACCTGTGCAACTCTTAATCTAGCCACATTTCTTCCAGAAAATGAAGAAAAAATAGA 241

Qy       3819 AYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGT 3878
              |:|||||||||||||:|||||  |||||||||| ||||||||||||||||||:||||:||
Db        242 ACATAACTGTCAACAAGTAATTGCTCAAACCTACGCCACTCGAGGGGACCTTCTAGAGGT 301

Qy       3879 TCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAG 3938
              |||:|||||||| |:|| |||||||||||  ||||||||||||  ||| |||||||||
Db        302 TCCCTTGACTGATTCCGA-CTTCAACTTGTACACTGATGGAAGTTTATTTTTAGAAAAAG 360

Qy       3939 GACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCC 3998
              |||||  |||||:|   |||||||||||||||||||||||||||:|||||||||||||||
Db        361 GACTTTAAAAAGCGTATTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCC 420

Qy       3999 TCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAAT 4058
              ||||||||||||||||||:||||:|||||| ||||||||:|||:|  ||  |||||||
Db        421 TCACTCCAGGAACTAGTGCTCAGCTGGCAGAAATAATAGCCCTCACTCAGGTGCTAGAAT 480

Qy       4059 TAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCC 4118
              |||||||| : |||||||:|||||||||||||||||:|:||||||:|||||||| |||
Db        481 TAGGAGAAAGGAAAAGGGTAAATATATATACAGACTCTAAATATGCTTACCTAGTCCTCC 540

Qy       4119 ATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCA 4178
              |||||||||::|||||||||:|:|||||||||||||||||||:|| :||  ||||||||
Db        541 ATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAAGGACAACCTATCA 600

Qy       4179 MACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAG 4238
              :|||||||||||||||||||:|||||||:||||:| ||||||| |||:|||||||||:|
Db        601 AACATCAGGAAGCCATTAGGAGATTATTATTGGCTGCACAGAAATCTAAAGAGGTGGCAA 660
```

```
Qy       4239 TCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGA 4289
              ||||||||| |:||||||||||| |||||||||:|||||||||||:|||:||
Db        661 TCTTACACTACTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGA 711


RESULT 236
BV667411/c
LOCUS       BV667411                 790 bp    DNA     linear   STS 16-APR-2005
DEFINITION  S217P6120FB9.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV667411
VERSION     BV667411.1  GI:62695408
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 790
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>790

  Query Match             8.1%;  Score 606.6;  DB 1;  Length 790;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  697;  Conservative    4;  Mismatches   83;  Indels    6;  Gaps    4;

Qy       1404 GGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASG 1463
              ||||| || |||  ||||||||     |||||| ||||||:|||||||||||| ||||| |
Db        788 GGAGAATGTTGCCACAGACATTGCTAAACTTGCGTGCTAGAAGGACTAAGGATAACTATG 729

Qy       1464 AAGAAARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATC 1523
              ||| | : || :||||   |||||||||||||| || |||||||| || ||||||||||
Db        728 AAG-AGGCCTGTGAATATTTCAATGATGTCCACTAT-ACACAGGGAAAAGGAAGAAAATC 671

Qy       1524 CTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTG 1583
              |||||||||||||||||||||||| |||||||||||||||| | ||||||||||||||||
Db        670 CTACTGCCTTTCTGGAGAGACTATGGGAGGCATTGAGGAAGCATACCTCTCTGTCACCTG 611

Qy       1584 ACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACA 1643
              |||||  |||||||||||||||||||  |||||||||||||| | ||||||||||||||
Db        610 ACTCTATTGAAGGCCAACTAATCTTAAAGGATAAGTTTATCACTTATTCAGCTGCAGACA 551

Qy       1644 TTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAA 1703
              |||||||| |||||||||| ||||| ||||||| |||||||||| ||||||||||||||||
Db        550 TTAGAAAAGACTTCAAAAGCCTGCCTTAGGCCCAGAGCAAAACTAAGAAACCCTATTGAA 491

Qy       1704 CTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACG 1763
              |||||||||:|||||||||||||||||||||||||||||||||||||| |||| ||||||||
Db        490 CTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCAGAACTGGACAAACA 431

Qy       1764 GGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCT 1823
              ||| ||| |||||||||||| |||||||||| |||||||||||  ||||||||||||||
Db        430 GGA-TAAGAAAAAGGCCACCACTTTAGTCATGGCCCTCAGGCAAGCAGACTTTGGAGGCT 372

Qy       1824 CTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTC 1883
              |||||| ||||||| |||||||||||||||||| ||||||||||||||||||||||| ||
Db        371 CTGGAACAGGGAAAGGCTGGGCAAATTGAATGCTTAATAGGGCTTGCTTCCAGTGAGGCT 312

Qy       1884 TACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCC 1943
              |||||  |||||||||| |||||||||  | ||||| |||||| |||| | |||||||||
Db        311 TACAATTACACTTTAAATAAGATTGTCTGAATAGAAATAAGCCACCCCCTTGTCCATGCC 252

Qy       1944 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGT 2003
              ||||||  |||||||||||||||||||||||||||||||||||||||  |||||| | |||||
Db        251 CCTTATATCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGATGAAGGTCCTCTGAGT 192

Qy       2004 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCAT 2063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        191 CAGAAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGTGCCAG 132

Qy       2064 CCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT- 2122
              |||||| |||||| |||| |||| || |||||||||||| |||||| |||||||| | |
Db        131 CCCATGACATCACTCTCATAGAGACCCGGGTATGCTTGATCATTGAAGGCCAGGAGGTTA 72

Qy       2123 --TGTCTCCTGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGT 2180
                |||| ||||||||||||  | | ||||| |||| |||||||||||| ||||||||||
Db         71 ACTGTCCCCTGGACACTGGCACAGCCTTCTCAGTCATACTCTTCTGTCCTGGACAACTGT 12

Qy       2181 CCTCCAGATC 2190
              ||||||| ||
Db         11 CCTCCAGTTC 2


RESULT 237
BV487214
LOCUS       BV487214                 725 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S216P6766FH7.T0 Masuku Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV487214
VERSION     BV487214.1  GI:62330884
```

```
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 725
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Masuku"
     STS             <1. .>725

  Query Match             8.1%;  Score 603.3;  DB 1;  Length 725;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  663;  Conservative    7;  Mismatches   54;  Indels    3;  Gaps    3;

Qy       6087 TACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAACTGTTGGATATG 6146
              ||| |||||||||| ||||||||   || ||||||||||||||||||||||||||||| ||
Db          1 TACTACCCTCACTAGGCTCCATCCCGTTTCGGCCCAAAACCCTACTAACTGTTGGATGTG 60

Qy       6147 CCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACTT 6206
              |||||||||| |||||:|||||   |||| ||||||||||||||||||||||||||||
Db         61 CCTCCCCCTGCACTTCAGGCCATA-CATTCATTCCCTGTACCTGAACAATGGAACAACTC 119

Qy       6207 CAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAAT 6266
              ||||||| ||||||||||||||||||||||||||||||||||||||||||:|||| ||
Db        120 CAGCACAAAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAGAT 179

Qy       6267 AACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTC 6326
              |||||||||||||||| |||||||||||||| |||||||| | ||||||||||||
Db        180 AACCCATACCTCAAACCTTACCTGTGTAAAATTTAACAATACTATAGACACAACCAACTC 239

Qy       6327 CCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATT 6386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATT 299

Qy       6387 TTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTT 6446
              |||||||||||||||||| |||||||||||||||| |||||||||||||||| |
Db        300 TTTTGTCTGTGGTACCTCAGACTATCGTTGTTTGAATGGCCCTTCAGAATCTATGTGCGT 359

Qy       6447 CCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGT 6506
              |||||||||||| ||  |||:|||:|||| |||||||||||||||||||| ||||||
Db        360 CCTCTCATTCTTACTG-TCCCTATGACCATTTACACTGAACAAGATTTATACAATTATGT 418

Qy       6507 CATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGT 6566
              || | |||||||| ||||||||||||||||||||||||||||||||| | |||||| ||||
Db        419 CACACCTAAGCCCGGCAACAAAAGAGTACCCATTCTTCCTTTTGTTGTTGGAGCAAGAGT 478

Qy       6567 GCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAA 6626
              ||||||   |||||||||||| |||||  || ||||||||||||||||||||||||||||
Db        479 GCTAGGCAGACTAGGTACTGACATTGATGGCATCACAACCTCTACTCAGTTCTACTACAA 538

Qy       6627 ACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCA 6686
              |||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db        539 ACTATCTCAAGAACTAAATGGTGACATGGAACAGGTCGCCGACTCCCTGGTCACCTTGCA 598

Qy       6687 AGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCT 6746
              ||||||| |||||||||||||||||||| |||||||:|||||||||||||||||||:|||
Db        599 AGATCAATTTAACTCCCTAGCAGCAGTGGTCCTTCAAAATCGAAGAGCTTTAGACTTGCT 658

Qy       6747 AACCGCTGARAGAGGGGGAA-CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTA 6805
              |||| | ||:|||||||||| |||||||||| |||||| |||||||||          ||
Db        659 AACCTCCGAAAGAGGGGGAACCCTGTTTATTTCTAGGGGGAGAATGCTGTCGGATATGTA 718

Qy       6806 ATCAATC 6812
              |||||||
Db        719 ATCAATC 725


RESULT 238
BV475768/c
LOCUS       BV475768                771 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P632397FF1.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV475768
VERSION     BV475768.1  GI:62235210
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
```

```
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 771
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>771

  Query Match             8.0%;  Score 598;  DB 1;  Length 771;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches  691;  Conservative   17;  Mismatches   62;  Indels   12;  Gaps    6;

Qy       2270 CTTTATTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGA 2329
             |||   |  |||||||||||  |   |||||||| |||||||||||| || |||  ||||
Db        771 CTTACTCTTTTTCACATGCTGTCTAAATTATGCCTGAAAGCCCCACTCCC-TGTAGGGGA 713

Qy       2330 GAGACATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCG 2389
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        712 GAGACATTCTAGCAAAAGCAGGGGCCACTATACACCTGAACATAGGAGAAGGAACACCCA 653

Qy       2390 TTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAA 2449
             |||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        652 TTTGTTGT-CCCCTGCTTGAGGAAGGAATTAATCCTGAAGCCTGGGCAACAGAAGGACAC 594

Qy       2450 TATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTT 2509
             |||||| ||| |||||||||| |||||||||| |||||||||||||||||| |||||
Db        593 TATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAATTAAACTAAAGGATTCCACCTCCTT 535

Qy       2510 TCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGT 2569
             |||||||||||||||||| |||| |||||| | ||||||||||||||||||||||||||
Db        534 TCCCTACCAAAGGCAGTACTCCCTTAGACCCGAAGCCCAACAAGGATTCCAAAAGATTGT 475

Qy       2570 TAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGT 2629
             ||| ||| |||||||||||||| |||||||||||||||  | | |||||| || | |
Db        474 TAAAGACCTAAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGTACCTGCAATACTCC--- 418

Qy       2630 AGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGG 2689
                 |||    |  | ||||||||||  |||||||||| ||||||||||||||||||||
Db        417 ----AATTTTAGGAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGATCTCAGG 363

Qy       2690 ATTATCAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMY 2749
             |||||||| |||||||||||| | |||| |||||| |||||| |||||||||| ||::
Db        362 ATTATCAAT-GAGGCCGTTGTCCCTCTATATCCAGCTATACCTAACCCTTATACTCTGCT 304

Qy       2750 TTCCCAAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTT 2809
             |||||||||| |||||||||||||||||||||||:|||||||||||:|||||||||| ||
Db        303 TTCCCAAATGCCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTT 244

Qy       2810 CTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCC 2869
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        243 CTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCC 184

Qy       2870 ARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATT 2929
             |: |||||||||| ||||||:|||||||||||||||||| |||||:|||||||||||
Db        183 AATGTCTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGAGATAGCCCCCATCTATT 124

Qy       2930 TGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCR 2989
             |||||||||| ||||||||||||||:| :|::|||||||||||| |||||||| || :
Db        123 TGGCCAGGCACTAGCCCAAGACTTGAGCCGGTTCTCATACCTGGAAACTCTTGTCTTTTG 64

Qy       2990 GTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCA 3049
             |||:|||||||||||||||||:||:|||::|||||||||||||| ||||||||||||||
Db          63 GTATGTGGATGATTTACTTTTAGCTGCCCATTCAGAAACCTTGTGACATCAAGCCACCCA 4

Qy       3050 AG 3051
              ||
Db           3 AG 2


RESULT 239
AY836231/c
LOCUS       AY836231                 717 bp    DNA     linear   PRI 20-SEP-2005
DEFINITION  Homo sapiens endogenous virus HERV-W clone W6 pol gene, partial
            sequence.
ACCESSION   AY836231
VERSION     AY836231.1  GI:56684755
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Forsman,A., Yun,Z., Hu,L., Uzhameckis,D., Jern,P. and Blomberg,J.
  TITLE     Development of broadly targeted human endogenous gammaretroviral
            pol-based real time PCRs: Quantitation of RNA expression in human
            tissues
  JOURNAL   J. Virol. Methods 129 (1), 16-30 (2005)
   PUBMED   15967513
REFERENCE   2  (bases 1 to 717)
  AUTHORS   Forsman,A., Yun,Z., Uzhameckis,D., Hu,L., Jern,P. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2004) Section of Virology, Dept of Medical
            Sciences, Uppsala University, Academic Hospital, Uppsala, Uppland
            751 85, Sweden
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="W6"
                     /note="endogenous virus: HERV-W"
     gene            complement(<1. .>717)
                     /gene="pol"
                     /note="3' end"
```

```
  Query Match             7.9%;  Score 592.7;  DB 1;  Length 717;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  626;  Conservative   51;  Mismatches   31;  Indels   19;  Gaps    4;

Qy       2700 GAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATA 2759
             ||||| ||||| | | |||| ||||||||||||||||||||||| ||::||||||||||
Db        711 GAGGCTGTTGTTCCTCTATAGCCAGCTGTACCTAGCCCTTATACTCTGCTTTCCCAAATA 652

Qy       2760 CCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCT 2819
             ||||||||||||||||||||||:||||||||||:|||||||||||||||||||||||||
Db        651 CCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTCAGGATGCCTTCTTCTGCATCCCT 592

Qy       2820 GTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAA 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        591 GTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCAACATCTCAA 532

Qy       2880 CTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCA 2939
             ||||||||||||:||||||||||||||||||||||:|||||||||||||||||||||||
Db        531 CTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTGGCCAGGCA 472

Qy       2940 TTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGAT 2999
             |||||||||||||||||:||:|::|||||||||||||  ||||||||||:|||:||||||
Db        471 TTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCTTCGGTAGGTGGAT 414

Qy       3000 GATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTM 3059
             ||||||||||:||:|||::|||||||||||||||||||||||||||||||:|||||:
Db        413 GATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTC 354

Qy       3060 AATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAG 3119
             ||||||||||:|||||||||||||:||||||||||:|:|:|||||:||||||||||||
Db        353 AATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCACAG 294

Qy       3120 CAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRY 3179
             |||||    ||||||||:||||:||||||||||||||||:||||||||||||||||:::
Db        293 CAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACAC 238

Qy       3180 ATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTT-CCT 3238
             ||||||||||||||||||||||||||||:|||||||||||||||||||:|::|| |||
Db        237 ATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGGGGATTCCCT 178

Qy       3239 TGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
             || |:|||:|||:|||||||||:|||||||||  |  : |  :||    ||||| |||
Db        177 TGNCGTAATAGGTTTCTGCCGAAAATGGATTCCCAGGTATGGCGAAATANCCCAGGNCAT 118

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||||:| ||| |||: :|||
Db        117 TAAATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTNAGANGGACNACTG 58

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             || :::||   | ||| |             ||||| ||| ||||||:||||||||||
Db         57 AANTAAAAAGTGGTTTTC-----------CAGGCCCTAACCCCAAGCCCCAGTGTTAAG 10

Qy       3419 YTTGCCA 3425
             :||||||
Db           9 TTTGCCA 3


RESULT 240
AB063619
LOCUS       AB063619                780 bp    mRNA    linear   PRI 26-DEC-2001
DEFINITION  Homo sapiens HERV-W pol element mRNA, isolate:HWP-FB12.
ACCESSION   AB063619
VERSION     AB063619.1  GI:18143409
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Isolation and Phylogeny of the Human Endogenous Retrovirus HERV-W
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 780)
  AUTHORS   Kim,H., Jeon,S. and Lee,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .780
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HWP-FB12"
                     /db_xref="taxon:9606"
                     /tissue_type="Fetal Brain"
     repeat_region   1. .780
                     /rpt_family="HERV-W pol element"

  Query Match             7.9%;  Score 589.4;  DB 1;  Length 780;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  673;  Conservative   64;  Mismatches   41;  Indels   16;  Gaps    9;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:| |||||:||||||||||||||| ||||||||||:||||:|||||||
Db          1 TCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||||||:|||||||||||||||:::| |||||| | |||||||||||||||||:|
Db         61 TAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||| ||:|||:|||:|||||| ||:|||||||||
Db        121 -AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATAACAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:||| :|||||||||:|||||:||||||||||||||||||||||||||||||
Db        180 -CCCAGGTATGGTGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACTCAGAAAG- 237

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||| |||| |:||||||:::||| ||||||||||||  || |
Db        238 CAATACCCATTTAGTAAGGTGGAAACCTGAAGCAGAAGCAGCTTTCCAGGCCCTAAAGAA 297

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||||||:||||||||||:||||||||:|||||||||||     :|||:
Db        298 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:  :||||||:||||||||||||| :|| :|| ||:||||||
Db        353 ATCACAG-AAAAAACAGGGGTAGCTCTAGGAGTCCTTACACATGTCTGAGAGACGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:|||| |||:|||||||||||||||||||:|:||||||
Db        412 GCAACCTGTGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| |||  ||| |||||||::|| ||||| |||||||||||| |||||| |
Db        472 TTTATGGGTAATGGCAGCAATAGCAGTCTTAATATCTAAAGCAGTTAAAATGATACAGTG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||     |||||||||:|||||::||||||||||||||||||||||
Db        532 AAGAGATCTTACTGTG-----ATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 586

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGT 3749
             |||||||||||||||||: |||||||||||:|||||||  ||||||||||:|||||||
Db        587 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCAGT 646

Qy       3750 GCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGA 3809
             |||||:|||||||||||||||||||||||||||:| ||||||||||||||||||||||
Db        647 GCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGA 706

Qy       3810 AAAGATARAAYATAACTGTCAACAAR-TAATTTCTCAAACCTATGCCAC-TCGAGGGGAC 3867
             |||||||:||:|||||||||||||: ||||| |||||||||||||| | ||||||||
Db        707 AAAGATAGAACATAACTGTCAACAAGTTAATTGCTCAAACCTATGCCACTTGGAGGGGAC 766
```

```
Qy       3868 CTTYTAGARGTTCC 3881
              |||:||| :|||||
Db        767 CTTTTAGGGGTTCC 780


RESULT 241
A46020
LOCUS       A46020                   741 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 9 from Patent EP0674004.
ACCESSION   A46020
VERSION     A46020.1  GI:2300292
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Mandrand,B., Perron,H., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and pathogenic and/or infections agent MSRV2 linked to
            multiple sclerosis, their nucleic components and their applications
  JOURNAL   Patent: EP 0674004-A 9 27-SEP-1995;
            BIO MERIEUX (FR)
COMMENT     Other publication DE 674004T 960919
            Other publication NO 953925 951204
            Other publication FI 954699 951003
            Other publication AU 1711495 950821
            Other publication CA 2141907 950805
            Other publication FR 2715939 950811
            Other publication FR 2715938 950811
            Other publication FR 2715937 950811
            Other publication FR 2715936 950811.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
              |||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:|||||||||:|     | |||||||
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   ||||||||||||||  ||  | | |||
Db         301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:||||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db         357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db         416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db         475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db         535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db         595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||: |||||||||| :|
Db         655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||||| ||
Db         715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 242
A57056
LOCUS       A57056                   741 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 9 from Patent WO9628552.
ACCESSION   A57056
VERSION     A57056.1  GI:3713039
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 9 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
              |||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:|||||||||:|     | |||||||
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   ||||||||||||||  ||  | | |||
```

```
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| |||||||||:|||||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| ||||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 243
A60121
LOCUS       A60121                   741 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 9 from Patent WO9706260.
ACCESSION   A60121
VERSION     A60121.1  GI:3715137
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 9 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db         301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db         357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db         416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| |||||||||:|||||||||||||||||||||:|:|||| |||||:||
Db         475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db         535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| ||||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db         595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db         655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db         715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 244
A79471
LOCUS       A79471                   741 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 9 from Patent EP0789077.
ACCESSION   A79471
VERSION     A79471.1  GI:6092479
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 9 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120
```

```
Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| |||||||||||:|||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| ||||||||||||||||||:|||||::||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: ||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740
```

```
RESULT 245
A80205
LOCUS       A80205                 741 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 9 from Patent WO9521256.
ACCESSION   A80205
VERSION     A80205.1  GI:6092937
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H. and Mallet,F.
  TITLE     MSRV1 VIRUS AND MSRV2 PATHOGENIC AND/OR INFECTIOUS AGENT ASSOCIATED
            WITH MULTIPLE SCLEROSIS, NUCLEIC ACID COMPONENTS AND APPLICATIONS
            OF SAME
  JOURNAL   Patent: WO 9521256-A 9 10-AUG-1995;
            BIO MERIEUX (FR); PERRON HERVE (FR)
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;
```

```
Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| |||||||||||:|||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: |||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| ||||||||||||||||||:|||||::||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:|||||||||||||||||: ||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740
```

```
RESULT 246
A80314
LOCUS       A80314                 741 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 9 from Patent EP0731168.
ACCESSION   A80314
VERSION     A80314.1  GI:6093041
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 9 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
```

```
                /db_xref="taxon:32644"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              |||||||:|||||||| ||:||||||||:|||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:||||:|||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||:: |||||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 247
AR035357
LOCUS       AR035357                741 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5871745.
ACCESSION   AR035357
VERSION     AR035357.1  GI:5952025
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Multiple sclerosis related virus
  JOURNAL   Patent: US 5871745-A 9 16-FEB-1999;

FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:|||||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              |||||||:|||||||| ||:||||||||:|||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:||||:|||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| |||||||:||||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||:: |||||||||||||||||||||||||||||||:|||||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 248
AR036198
LOCUS       AR036198                741 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5871996.
ACCESSION   AR036198
VERSION     AR036198.1  GI:5952866
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
```

```
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Retrovirus agents MSRV1 and MSRV2 associated with multiple
            sclerosis
  JOURNAL   Patent: US 5871996-A 9 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||| :|||||:|||||||||| |:|||||||||||||:||||||||||:|
Db             1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db            61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db           121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db           181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:||||||||||||||||||||||||||||||
Db           241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|   |||||||||||||   ||  | | |||
Db           301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: |||||
Db           357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:|||||||||||||||||:|||:|||||:||||||||||||:
Db           416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| ||||||||:|||||||||||||||||:|:|||| |||||:||
Db           475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db           535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db           595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db           655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db           715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 249
AR036718
LOCUS       AR036718                741 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5800980.
ACCESSION   AR036718
VERSION     AR036718.1  GI:5954574
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Detection of MSRV1 virus and MSRV2 pathogen and/or infective agent
            associated with multiple sclerosis, by nucleic acid hybridization
  JOURNAL   Patent: US 5800980-A 9 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||| :|||||:|||||||||| |:|||||||||||||:||||||||||:|
Db             1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db            61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db           121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             |||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db           181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:||||||||||||||||||||||||||||||
Db           241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|   |||||||||||||   ||  | | |||
Db           301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||| ||:|||||||:|||||||| ||||:||:|||:: |||||
Db           357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:|||||||||||||||||:|||:|||||:||||||||||||:
Db           416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| ||||||||:|||||||||||||||||:|:|||| |||||:||
Db           475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| |||||||||||:: |||||||||||||||||||||||||||||:||||||
Db           535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db           595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db           655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db           715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 250
AR076980
LOCUS       AR076980                741 bp    DNA     linear   PAT 31-AUG-2000
```

```
DEFINITION  Sequence 9 from patent US 5962217.
ACCESSION   AR076980
VERSION     AR076980.1  GI:10003726
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and MSRV2 pathogen and/or infective agent associated
            with multiple sclerosis, and biopolymer constituents thereof
  JOURNAL   Patent: US 5962217-A 9 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||||| || ||| |||||||||: |||:|||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||||  ||:||||||||:|||||||  ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:||||||||||||||:|||:|||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| |||||||:|||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| ||||||||||:: |||||||||||||||||||||||||||||||:|||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||| || 
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740
```

```
RESULT 251
AR094701
LOCUS       AR094701                741 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 9 from patent US 6001987.
ACCESSION   AR094701
VERSION     AR094701.1  GI:10021867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 9 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:| 
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
             :|:|||||:|||||||||||||| || ||| |||||||||: |||:|||||||||||||| 
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
             |:|| ||||||||||||:::|||||||||||||||| ||||||||||:|||||||||| 
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
             ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | ||||||| 
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||| 
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
             |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | ||| 
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
             ||||||||:|||||||||  ||:||||||||:|||||||  ||||:||:|||:: ||||| 
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
             |||||||||:||: |||||:||||||||||||||:|||:|||||:||||||||||||: 
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
             :||||::|||||| |||||||:|||||||||||||||||:|:|||| |||||:|| 
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
             ||| ||| ||||||||||:: |||||||||||||||||||||||||||||||:|||| 
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||| 
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :| 
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714
```

```
Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
                ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 252
AR183936
LOCUS       AR183936                741 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 9 from patent US 6342383.
ACCESSION   AR183936
VERSION     AR183936.1  GI:20227905
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Cell lines and viral isolates associated with multiple sclerosis
  JOURNAL   Patent: US 6342383-A 9 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
                |||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
                :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
                |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
                ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
                :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||||
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
                |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db         301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
                ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db         357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
                |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db         416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
                :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db         475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
                ||| ||| ||||||||||:: |||||||||||||||||||||||||||||:||||||
Db         535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
                ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||
Db         595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
                ||||||||||: ||||||||| |:|||||||||||||||||: |||||||||| :|
Db         655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
                ||| |||||||||||||||||| ||
Db         715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 253
AR343148
LOCUS       AR343148                741 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 9 from patent US 6579526.
ACCESSION   AR343148
VERSION     AR343148.1  GI:33738650
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 9 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
                |||||||||||||| :|||||:|||||||| |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
                :|:|||||:|||||||||||| || ||| |||||||||: |||:||||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
                |:|| |||||||||||||:::|||||||||||||||| ||||||||||:|||||||||||
Db         121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
                ||||||||||:|::|||||| ||:| | :|||:||||||||||:|     | |||||||
Db         181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
                :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||||
Db         241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
                |||||:||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db         301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
                ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:|||:: |||||
Db         357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
                |||||||||:||: |||||:|||||||||||||||:|||:|||||:||||||||||||:
Db         416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
                :||||::|||||| |||||||:|||||||||||||||||||:|:|||| |||||:||
Db         475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
                ||| ||| ||||||||||:: |||||||||||||||||||||||||||||:||||||
```

```
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
             ||| |||||||||||||||||:|||||::|||||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
             ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 254
AR344305
LOCUS       AR344305                741 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 9 from patent US 6582703.
ACCESSION   AR344305
VERSION     AR344305.1  GI:33740246
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 9 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             7.8%;  Score 586;  DB 1;  Length 741;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  628;  Conservative   62;  Mismatches   48;  Indels    8;  Gaps    5;

Qy       3039 CAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSA 3098
             |||||||||||||  :|||||:|||||||||  |:|||||||||||||:||||||||||:|
Db          1 CAAGCCACCCAAGAACTCTTAAATTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCA 60

Qy       3099 RARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACC 3158
              :|:|||||:|||||||||||||  || ||| |||||||||: |||:|||||||||||||||
Db         61 AAGGCTCAGCTCTGCTCACAGGAGATTAGATACTTAGGGTTAAAATTATCCAAAGGCACC 120

Qy       3159 ARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTA 3218
              |:|| ||||||||||||:::|||||||||||||||| ||||||||||:||||||||||
Db        121 AGGGGCCTCAGTGAGGAACGTATCCAGCCTATACTGGGTTATCCTCATCCCAAAACCCTA 180

Qy       3219 AAGCAACTAAGRGRRTTCCTTGGCRT-AAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||:|::|||||| ||:| | :|||:||||||||:|     | |||||||
Db        181 AAGCAACTAAGAGGGTTCCTTAGCATGATCAGGTTTCTGCCGAAAACAAGATTCCCAGGT 240

Qy       3278 W-TGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACC 3336
              :    |:|||||||||| :|||||:|||||:|||||||||||||||||||||||||||||
Db        241 ACAACCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACC 300

Qy       3337 CATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTT 3396
              |||||:|||||||||:|:|| ||   :|   |||||||||||||  ||  | | |||
Db        301 TATTTAGTAAGATGGACACCTAAACAGAA---GGCTTTCCAGGCCCTAAAGAAGGCCCT- 356

Qy       3397 AAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAG 3456
              ||||||||:|||||||||  ||:||||||||:|||||||| ||||:||:|||:: |||||
Db        357 -AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTATATGGCACAG 415

Qy       3457 AAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCY 3516
              |||||||||:||: |||||:|||||||||||||||:|||:|||||:|||||||||||||:
Db        416 -AAAAAACAGGAATCGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTGCAACCC 474

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||:|:|||| |||||:||
Db        475 GTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATNGTTTATGG 534

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||| ||| ||||||||||||:: ||||||||||||||||||||||||||||:||||||
Db        535 GTAATGGNGGCAGTAGCAGTCTNAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGAT 594

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              ||| |||||||||||||||||:||||||::|||||||||||||||||||||||||||||
Db        595 CTTNCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGG 654

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||: ||||||||| |:||||||||||||||||||: |||||||||| :|
Db        655 TTGTCAGACAACCATTTACTTAANTATCAGGCTCTATTACTTGAAGAGCCAGTGCTGNGA 714

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCC 3782
              ||| |||||||||||||||||| ||
Db        715 CTGCGCACTTGTGCAACTCTTAAACC 740


RESULT 255
BV527275
LOCUS       BV527275                778 bp    DNA     linear   STS 08-APR-2005
DEFINITION  G591P610727RD4.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV527275
VERSION     BV527275.1  GI:62405059
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 778)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 778
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
```

consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .778
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>778

Query Match 7.8%; Score 585.501; DB 1; Length 778;
Best Local Similarity 82.7%; Pred. No. 0;
Matches 645; Conservative 41; Mismatches 89; Indels 5; Gaps 3;

```
Qy       3728 CTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCN 3787
Db          1 CTCTATTACTTAAAGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCCA 60

Qy       3788 CATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAA 3847
Db         61 CATTTCTTTCAGACAATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAA 120

Qy       3848 CCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTG 3907
Db         121 CCTATGGCGCTAGAGGGGACCTTTTAGAGTTTCCCTTGACTGATCCTGACC-TCAACTTG 179

Qy       3908 TATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTC 3967
Db         180 TATACTGATTGAAGTCCTTTTTCTACAAAGAAACTTCTAAAAGTGAGGTATGCAGTGGTC 239

Qy       3968 AGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCA 4027
Db         240 AGTGATAATGGAATACTTAAAAGTAATCCCCTCACTCCAAGAACTAGTGCTCAGCTGGCA 299

Qy       4028 GAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATAT 4087
Db         300 GAACTAATAGCCCTCACTTGGGCACTAGAATTAGGAGAAAGAAAAAAGGTAAATATATAT 359

Qy       4088 ACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGG 4147
Db         360 ACAGCCTCTAAGTATGCTTACCTACTCCTCCATGCCCATGCAACAATGTAGAGAGAAAAG 419

Qy       4148 GAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGARATTATTA 4207
Db         420 GAATTCCTAACTTCCGAAGGAACATGTATCAAACATCAGGAAGCCATTAGGAGATAATTA 479

Qy       4208 YTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGA 4267
Db         480 TTGGCTGTACAGAAACCTAAAAAGGTAACAATCTTACACTGCCAAGGTCACCAGAAAGGA 539

Qy       4268 AAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGG 4327
Db         540 AAAGAAAGGAAAATAGAAAGGAACCACCAAGCGGATACTCAAGTCAAAAGAGCCGCAAGG 599

Qy       4328 CAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGG 4387
Db         600 CAGGACCCTCCATTAGAAATGCTTATAGAAGGACCCCTAGTATAGGGTAATCCCCTCCAG 659

Qy       4388 GAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAAC---CTCACGAGGCAGT 4444
Db         660 GACACCAAGCCCCAGCATTCAGCAGGAAAAATAGAATAGGGAACCCTCACAGGTCATACT 719
```

```
Qy       4445 TTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATC 4504
Db         720 TTCCTCCCCCCTCAGATGCTAGCCACTGAAGAA-AGAAAATACTTTTGTCTGCAGCTAAC 778
```

RESULT 256
BV479791
LOCUS BV479791 661 bp DNA linear STS 07-APR-2005
DEFINITION qcf65g06.g1 Clint Pan troglodytes verus STS genomic, sequence
tagged site.
ACCESSION BV479791
VERSION BV479791.1 GI:62323152
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 661)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 661
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.
FEATURES Location/Qualifiers
source 1. .661
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"

```
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>661

  Query Match             7.7%;  Score 579.9;  DB 1;  Length 661;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  609;  Conservative   10;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       6271 CATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAA 6330
             |||||||||||||| |||||||||||||||| |||||||| | ||||||||||||||||||
Db             1 CATACCTCAAACCTTACCTGTGTAAAATTTAACAATACTATAGACACAACCAACTCCCAA 60

Qy       6331 TGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTT 6390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTT 120

Qy       6391 GTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTC 6450
             |||||||||||||||| |||||||||||||||||| |||||||||||||||| |||||
Db           121 GTCTGTGGTACCTCAGACTATCGTTGTTTGAATGGCCCTTCAGAATCTATGTGCGTCCTC 180

Qy       6451 TCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATA 6510
             ||||||||| ||  |||:|||:|||| |||||||||||||||||||||| |||||||| |
Db           181 TCATTCTTACTG-TCCCTATGACCATTTACACTGAACAAGATTTATACAATTATGTCACA 239

Qy       6511 TCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTA 6570
              ||| |||| |||||||||||||||||||||||||||| | ||||||||||||||||||
Db           240 CCTACGCCCGGCAACAAAAGAGTACCCATTCTTCCTTTTGTTGTTGGAGCAGGAGTGCTA 299

Qy       6571 GGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTA 6630
              ||   |||||||||||| |||||  || ||||||||||||||||||||||||||||||||
Db           300 GGCAGACTAGGTACTGACATTGATGGCATCACAACCTCTACTCAGTTCTACTACAAACTA 359

Qy       6631 TCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGAT 6690
             |||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||
Db           360 TCTCAAGAACTAAATGGTGACATGGAACAGGTCGCCGACTCCCTGGTCACCTTGCAAGAT 419

Qy       6691 CAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACC 6750
             ||| |||||||||||||||| |||||||:|||||||||||||||||||:|||||||
Db           420 CAATTTAACTCCCTAGCAGCAGTGGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACC 479

Qy       6751 GCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 6810
              | ||:||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db           480 TCCGAAAGAGGGGAAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAA 539

Qy       6811 TCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGAR 6870
             |||| ||||||||| |||||||||:|| ||||||:||||| |||||||| ||:||||||:
Db           540 TCCGAAATCGTCACCGAGAAAGTTAAAAAAATTCGAGATCAAATACAACATAGAGCAGAG 599

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| |||||| |||||||||||| |||||| ||:||||||||||||||||||||||
Db           600 GAGCTTCAAAACACCGGACCCTGGGGCTTCCTCACCCAATGGATGCCCTGGATTCTCCCT 659

Qy       6931 T 6931
              |
Db           660 T 660


RESULT 257
BV470750/c
LOCUS       BV470750                757 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P609728FG2.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV470750
VERSION     BV470750.1  GI:62230178
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 757
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>757

  Query Match             7.7%;  Score 579.8;  DB 1;  Length 757;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  619;  Conservative   62;  Mismatches   39;  Indels    6;  Gaps    5;

Qy       2804 CTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTC 2863
              |||  ||||||| |||||||||||||| |||||||||||||||||||||||||||| ||
Db        720 CTTTGTCTGCATTCCTGTACATCCTGAGTCTCAATTCTTGTTTGCCTTTGAAGATCCTGT 661

Qy       2864 AAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCA 2923
              ||| ||:  ||||| |||||||||||:||||||| |||||||||||||||||:|| ||
Db        660 GAACTCAATGTCTCACCTCACCTGGACTGTTTTACCGCAAGGGTTCAGGGATAGCCCACA 601

Qy       2924 TCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGT 2983
              |||||||||||||||||||||||||||||||:||:|::|||||||||||||||||||
Db        600 TCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAATTCTCATACCTGGACACTCTTGT 541

Qy       2984 CCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGC 3043
```

```
Db        540 CCTTCAGTAGGTGGATGATTTACTTTTAGCCACCCGTTCAGAAACCTTGTGCCATCAAGC 481

Qy       3044 CACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGC 3103
Db        480 CACCCAAGTGCTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGC 421

Qy       3104 TCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGC 3163
Db        420 TCAGCTCTTCTCACAGCAGGTTAAATACTTAGGGCTACAGTTATCCCAAGGCACCAGGGC 361

Qy       3164 CCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCA 3223
Db        360 CCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCA 301

Qy       3224 ACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCR 3283
Db        300 ACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCAAATATGGATT-CCCAAGTACAGCA 242

Qy       3284 AAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAR 3343
Db        241 AAATAGCCAGACCATTATATACACCAATTAAGGAAACTC--AAAAGCCAATACCATTTAG 184

Qy       3344 TAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCA 3403
Db        183 TAAGATGGACACCTGAAGCAGAAGCAGCTTTCCAGG-CACTAAAGAAGGCCTT-AACCCA 126

Qy       3404 AGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAA 3463
Db        125 AGCCCCAGTGTTAAGTTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-GAAAAA 67

Qy       3464 CAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRY 3523
Db          66 CAGGAAGAGCTCTAGGAGTCTTTACACAGGTCCGAGGGACGAGCTTGCAACCCATGCCAC 7

Qy       3524 ACCTGA 3529
Db           6 ACCTGA 1


RESULT 258
AY836232/c
LOCUS       AY836232                705 bp    DNA     linear   PRI 20-SEP-2005
DEFINITION  Homo sapiens endogenous virus HERV-W clone W4 pol gene, partial
            sequence.
ACCESSION   AY836232
VERSION     AY836232.1  GI:56684756
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   Forsman,A., Yun,Z., Hu,L., Uzhameckis,D., Jern,P. and Blomberg,J.
  TITLE     Development of broadly targeted human endogenous gammaretroviral
            pol-based real time PCRs: Quantitation of RNA expression in human
            tissues
  JOURNAL   J. Virol. Methods 129 (1), 16-30 (2005)
   PUBMED   15967513
REFERENCE   2  (bases 1 to 705)
  AUTHORS   Forsman,A., Yun,Z., Uzhameckis,D., Hu,L., Jern,P. and Blomberg,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2004) Section of Virology, Dept of Medical
            Sciences, Uppsala University, Academic Hospital, Uppsala, Uppland
            751 85, Sweden
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="W4"
                     /note="endogenous_virus: HERV-W"
     gene            complement(<1. .>705)
                     /gene="pol"
                     /note="3' end"

  Query Match             7.7%;  Score 578.9;  DB 1;  Length 705;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  613;  Conservative   50;  Mismatches   24;  Indels    9;  Gaps    5;

Qy       2700 GAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATA 2759
Db        699 GAGGCTGTTGTTCCTCTATAGCCAGCTGTACCTAGCCCTTATACTCTGCTTTCCCAAATA 640

Qy       2760 CCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCT 2819
Db        639 CCAGAGGAAGCAGAGTGGTTTACGGTCCTGGACCTTCAGGATGCCTTCTTCTGCATCCCT 580

Qy       2820 GTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAA 2879
Db        579 GTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCAACATCTCAA 520

Qy       2880 CTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCA 2939
Db        519 CTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTGGCCAGGCA 460

Qy       2940 TTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGAT 2999
Db        459 TTAGCCCAAGACTTGAGCCAATCCTCATACCTGGACA--CTTGTCCTTTGGTAGGTGGAT 402

Qy       3000 GATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTM 3059
Db        401 GATTTACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTC 342

Qy       3060 AATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAG 3119
Db        341 AATTTCCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCACAG 282

Qy       3120 CAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRY 3179
Db        281 CAGGT----TACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACAC 226

Qy       3180 ATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT 3239
Db        225 ATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAACTAAGGGGATTCCTT 166

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
Db        165 GGCGTAATANGTTTCTGCCGAAAATGGATT-NCCAGGTATGGCGAAATANCCAGGTCATT 107

Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMC-TG 3358
Db        106 AAATACACTAATTAAGGAAA-TCANAAAGCCAATACCCATTTAGTAANANGGACNACTTG 48

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCC 3394
Db          47 AAGTAAAAATGGGTTTTCAGGGCCCTTAACCNAGCC 12


RESULT 259
BV477069
LOCUS       BV477069                787 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P638207FC7.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV477069
VERSION     BV477069.1  GI:62236511
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
```

```
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 787
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .787
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>787

  Query Match             7.7%;  Score 577.9;  DB 1;  Length 787;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  675;  Conservative   17;  Mismatches   68;  Indels   13;  Gaps    6;

Qy       2279 TTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTC 2338
              |||  ||||||| |||||||||| ||||   || ||| ||||||||||||||||||||
Db          3 TTTTCCATGCTTGTCTAATTATGCCTGAA--GCCTACTCCCTTGTTAGGGAGAGACATTC 60

Qy       2339 TAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTN 2398
              |||||||||||||||||| |||||||||||||||||||||||||||||| ||||||||
Db         61 TAGCAAAAGCAGGGGCCACTATACACCTGAACATAGGAGAAGGAACACCCATTTGTTGT- 119

Qy       2399 CCCCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGA 2458
              |||||||||||||||||||||||||||||| |||||||||||||||| |||||| ||
Db        120 CCCCTGCTTGAGGAAGGAATTAATCCTGAAGCCTGGGCAACAGAAGGACACTATGGATGA 179

Qy       2459 GCCAAAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCA 2518
              | |||||||||| |||||||||| |||||||||||||||||| ||||||||||||||
Db        180 G-CAAAGAATGCCTGTCCTGTTCAAATTAAACTAAAGGATTCCACCTCCTTTCCCTACCA 238

Qy       2519 AAGGCAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTT 2578
              |||||||||| |||| |||||| | |||||||||||||||||||||||||||||| ||| |
Db        239 AAGGCAGTACTCCCTTAGACCCGAAGCCCAACAAGGATTCCAAAAGATTGTTAAAGACCT 298

Qy       2579 AAAAGCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTG 2638
              |||||||||||||| ||||||||||||||   | | |||||| || | |         |||
Db        299 AAAAGCCCAAGGCCTAGTAAAACCATGCAGTAGTACCTGCAATACTCC-------AATTT 351

Qy       2639 AGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAAT 2698
                |  | ||||||||||  |||||||||| ||||||||||||||||||||||||||||||
Db        352 TAGGAGTACAGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTATCAAT 410

Qy       2699 GGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAAT 2758
              |||||||||||| | |||| |||||| |||||| |||||||||| ||::|||||||||
Db        411 -GAGGCCGTTGTCCCTCTATATCCAGCTATACCTAACCCTTATACTCTGCTTTCCCAAAT 469

Qy       2759 ACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCC 2818
              |||||||||||||||||||||:||||||||||:|||||||||| ||||||||||
Db        470 GCCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCC 529

Qy       2819 TGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCA 2878
              |||||||||||||||||||||||||||||||||||||||| ||||||||||:  |||||
Db        530 TGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCA 589

Qy       2879 ACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGC 2938
              |||||| ||||||:|||||||||||||||||||| |||||:|||||||||||||||||||
Db        590 ACTCACTTGGACTGTTTTACCCCAAGGGTTCAGAGATAGCCCCCATCTATTTGGCCAGGC 649

Qy       2939 ATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGA 2998
              | |||||||||||||||:| :|::|||||||||||| |    |||| || :|||:|||||
Db        650 ACTAGCCCAAGACTTGAGCCGGTTCTCATACCTGGAAAACTCTGTCTTTTGGTATGTGGA 709

Qy       2999 TGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAG 3051
              ||||||   |||:||:|||::|||||||||||||||||    |||||||| ||
Db        710 TGATTTTACTTTAGCTGCCCATTCAGAAACCTTGTGCCCATCAGCCACCCCAG 762


RESULT 260
BV647045
LOCUS       BV647045                792 bp    DNA     linear   STS 16-APR-2005
DEFINITION  S215P6949RB5.T0 Clara Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV647045
VERSION     BV647045.1  GI:62675015
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 792
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
```

including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .792
/organism="Pan troglodytes troglodytes"
/mol_type="genomic DNA"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/clone_lib="Clara"
STS <1. .>792

Query Match 7.7%; Score 575.9; DB 1; Length 792;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 692; Conservative 5; Mismatches 91; Indels 9; Gaps 6;

```
Qy       1414 GCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCT 1473
             | | |||||||||||||||||| ||||||:||||||||  |||||||:| |   |: ||
Db          1 GATTCAGACATTTGCTAACTTGCGTGCTAGAAGGACTAACAAAAACTAGGCATT-AGCCT 59

Qy       1474 AYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTT 1533
              :|| ||| |||||||||||||| |||||||||  || ||||||||||||||||||||
Db         60 GTGAGTTATTCAATGATGTCCACTATAACACAGGTTAA-GGAAGAAAATCCTACTGCCTT 118

Qy       1534 TCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGA 1593
             |||||| || |||||||||||||| ||||||| | ||||||||||||||||||   |||
Db        119 TCTGGACAGGCTAAGGGAGGCATTAAGGAAGCATACCTCTCTGTCACCTGACTCTATTGA 178

Qy       1594 AGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACA-TTAGAAAAA 1652
             |||||||||||| ||||||  |||||||||||||||||||||||||||| ||| |||||
Db        179 AGGCCAACTAATTTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGACATTTAAAAAAA 238

Qy       1653 ACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAAC 1712
             |||||||||||||||| |||||| |||||||||||||| ||||||||||||||||||||
Db        239 ACTTCAAAAGTCTGCCTTAGGCCTGGAGCAAAACTTAAAAACCCTATTGAACTTGGCAAC 298

Qy       1713 YTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAA 1772
              :| ||||||||||||||||||||||||||||||||||||| |||||||||||||| || ||| |
Db        299 CTTGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAGGACAAACGAGA-TAAGA 357

Qy       1773 AAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAG 1832
             |||||||||| |||||||||||| |||||||||||  |||||||||||||||||||| |
Db        358 AAAAGGCCACTGCTTTAGTCATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGGAACAT 417

Qy       1833 GGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGAC 1892
             |||||  ||| |||||| |||||| ||  |||||||||||| ||| | ||||||||||||
```

```
Db        418 GGAAAGCCTGAGCAAATCGAATGCTTA--AGGGCTTGCTTCTAGTACAGTCTACAAGGAC 475

Qy       1893 ACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTC 1952
             |||||||||||||||| | | ||| | |||||| ||   |||||||||||||||| ||
Db        476 ACTTTAAAAAAGATTGTTCGAATAGCAATAAGCCACCGTCTCGTCCATGCCCCTTATGTC 535

Qy       1953 AAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCA 2012
             |||||||||| | | ||||||||| |||||||||  ||| || | ||||||||||||||
Db        536 AAGGGAATCAGTTGGAGGCCCACTTCCCCAGGGGATGAAGTTCCTCTGAGTCAGAAGCCA 595

Qy       2013 CTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCA 2072
             |||||||||||||||||| |||||||||||||| || ||||| |||| |||||||||
Db        596 CTAACCAGATGATCCAGCAGTAGGACTGAGGGTGCCCGGTGCAAGTGCCAGCCCATGCCA 655

Qy       2073 TCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCC 2129
             |||||||||||||| ||  |||||||||||||||||||||||||||| | |   |||||||
Db        656 TCACCCTCACAGAGACCCAGGTATGCTTGACCATTGAGGGCCAGGAGGTTAACTGTCTCC 715

Qy       2130 TGGACACTGGTGCGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGAT 2189
             |||||||||  | | ||||| |||||||||| |||||||||||||||||||  |  ||
Db        716 TGGACACTGGCTCAGCCTTCTCAGTCTTACTCTCCTGTCCCGGACAACTGTCCCTCCAAT 775

Qy       2190 CTGTCACTATTCTGAGG 2206
             || ||||||| |   ||
Db        776 CTATCACTATCCAAGGG 792
```

RESULT 261
BV493485
LOCUS BV493485 781 bp DNA linear STS 07-APR-2005
DEFINITION S217P6797FD3.T0 Noemie Pan troglodytes troglodytes STS genomic,
sequence tagged site.
ACCESSION BV493485
VERSION BV493485.1 GI:62337155
KEYWORDS STS.
SOURCE Pan troglodytes troglodytes
ORGANISM Pan troglodytes troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 781)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 781
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the

base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .781
/organism="Pan troglodytes troglodytes"
/mol_type="genomic DNA"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/clone_lib="Noemie"
STS <1. .>781

Query Match 7.6%; Score 572.8; DB 1; Length 781;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 652; Conservative 34; Mismatches 78; Indels 8; Gaps 5;

```
Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
Db          7 AGTATTGCTCAACTACGCCACTCGAGGGGACTTCTT--TTGTTCCCTTGACTGATTCCGA 64

Qy       3897 CCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGT 3956
Db         65 -CTTCAACTTGTACACTGATGGAAGTTTATTTTTAGAAAAAGGACTTTAAAAAGCGTATT 123

Qy       3957 ATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 4016
Db        124 ATGCAGTGGTCAGTGATAATGGAATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTG 183

Qy       4017 CTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGG 4076
Db        184 CTCAGCTGGCAGAAATAATAGCCCTCACTCAGGTGCTAGAATTAGGAGAAAGGAAAAGGG 243

Qy       4077 YAAATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATAT 4136
Db        244 TAAATATATATACAGACTCTAAATACGCTTACCTAGTCCTCCATGCCCATGCAGCAATAT 303

Qy       4137 GSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTA 4196
Db        304 GGAGAGAAAGGGAATTCCTAACTTCTGAAGGACAACCTATCAAACATCAGGAAGCCATTA 363

Qy       4197 GGARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTC 4256
Db        364 GGAGATTATTATTGGCTGCACAGAAATCTAAAGAGGTGGCAATCTTACACTACTGGGGTC 423

Qy       4257 ATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAA 4316
Db        424 ATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCGCCAAGTGGATATTGAAGCCAAAA 483

Qy       4317 GAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTA 4376
Db        484 GAGCCACAAGGCAGGACCCTCCATAAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTA 543

Qy       4377 ATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCAC 4436
Db        544 ATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCAAGAGAGAGAGAATGGGGAACCTCAC 603

Qy       4437 GAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGC 4493
Db        604 AAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCCACTGAAGAAGGGAAAATACTTTTGT 663
```

```
Qy       4494 CTGCAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCG 4553
Db        664 CTGCAGCTAACCCATGGAAATTA-TTTAAACCCTTCACCAAACCTTTCACTTAGGCATTG 722

Qy       4554 ATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAA 4605
Db        723 ATAGCACCCATCAGATGGCCAAAATTATA-TTACTGTACCAGCCCTTTCAAA 773
```

RESULT 262
BV650860/c
LOCUS BV650860 783 bp DNA linear STS 16-APR-2005
DEFINITION S217P62327RH2.T0 Noemie Pan troglodytes troglodytes STS genomic, sequence tagged site.
ACCESSION BV650860
VERSION BV650860.1 GI:62678830
KEYWORDS STS.
SOURCE Pan troglodytes troglodytes
ORGANISM Pan troglodytes troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 783)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 783
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.

```
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>783

  Query Match             7.6%;  Score 569.4;  DB 1;  Length 783;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  655;  Conservative   36;  Mismatches   64;  Indels   20;  Gaps    6;

Qy       3900 TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATG 3959
             |||||||||||||||||| | |    |||||||||||||||||| |||||:||||||||
Db        776 TCAACTTGTATACTGATGAAGTT---CTTGTAGAAAAAGGACTTC-AAAAGTGGGGTATG 721

Qy       3960 CAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTY 4019
             ||||||||||||||||||||:|  ||||||| | ||||||| ||||||||||||||:
Db        720 CAGTGGTCAGTGATAATGGAATACTGCAAAGTAACCTCCTCACT-CAGGAACTAGTGCTC 662

Qy       4020 AGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA 4079
             ||||:|||||||||| ||||:|||:|:|||||||| |||| |||||||:|||||||:||
Db        661 AGCTGGCAGAACTAACAGCCCTCACTTGGGCACTAAAATTGGGAGAAGAAAAAAGGGTAA 602

Qy       4080 ATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSA 4139
             |||||||||||||||||:|: |||||:|||||||| |||||||||  |::|||||||:|
Db        601 ATATATATACAGACTCTAAGCATGCTTACCTAGTCCTCCATGCCCACGCAGCAATATGGA 542

Qy       4140 RAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGGA 4199
             :|||||||||||  | |||||||:|||:||           ||||||||||| |||||||
Db        541 GAGAAAGGGAATGCTTAACTTCTGAGGGA-----------ACATCAGGAAGCTATTAGGA 493

Qy       4200 RATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCATC 4259
             :|||||||:||||:|||||||||||||:||||||||:|||||||||||||:|||||||||
Db        492 GATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCTGGGGTCATC 433

Qy       4260 ANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGAG 4319
             | |||||||||:||||||||||:|||:|||:  |||||| :||||||||||:|||||||
Db        432 AGAAAGGAAAGGAAAGGGAAATAGAAGGGAACCACCAAGCGGATATTGAAGCCAAAAGAG 373

Qy       4320 CTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGGTAATC 4379
             | |||||||      |  |||||||||||||  ||   ||| |||||| ||||||||
Db        372 CCGCAAGGC-TCTTTAGCATTTAGAAATGCTTATAGAAGGACCCCTAGTATGGGGTAATC 314

Qy       4380 CCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAATGGGGAACCTCACGAG 4439
             || || ||||||||||||||||||||||||||||||||||||  |||||||| ||||||||||
Db        313 CCCTCTGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAATCGAATGGGGTACCTCACGAG 254

Qy       4440 G---CAGTTTTCTCCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAATACTTTTGCCTG 4496
             |    ||||| |||||||| ||| || ||||||||||||||| ||||||||||||||||
Db        253 GACATAGTTTCCTCCCCTCAGGATGGCTAGCCACTGAAGAAGGAAAAATACTTTTGCCTG 194

Qy       4497 CAACTATCCAATGGAAATTACTTAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATA 4556
             || ||| |||||||||||||||||||||||||||| ||||||||||||||||||||| |||
Db        193 CAGCTAACCAATGGAAATTACTTAAAACCCTTCACCAAACCTTTCACTTAGGCATTGATC 134

Qy       4557 GCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 4616
             ||||||||||:|  |||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GCACCCATCAGACAGCCAAATTATTATTTACTGGACCAGGCCTTTTCAAAACTATCAAGC 74

Qy       4617 ARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYATCGCCA 4671
             |:|||:|||||||||||||| |||| ||:|:|||||||||||||  :    ||||
Db          73 AGATAGTCAGGGCCTGTGAAATGTGACAAAGAAATAATCCCCTGCACTCAGGCCA 19


RESULT 263
BD136202
LOCUS       BD136202                758 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION   BD136202
VERSION     BD136202.1  GI:23231147
KEYWORDS    JP 2002509437-A/12.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
            Perron,H. and Mandrand,B.
  TITLE     Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: JP 2002509437-A 12 26-MAR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002509437-A/12
            PD   26-MAR-2002
            PF   07-JUL-1998 JP 1999508255
            PR   07-JUL-1997 FR     97/08816
            PI   GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC  PI
              BEDIN,
            PI   MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,
            PI   BERNARD MANDRAND
            PC   C12N15/48,C12Q1/70,C07K14/15,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Retroviral nucleic material and nucleotide fragments, in  CC
                 particular,
            CC   associated with multiple sclerosis and/or
            rheumatoid arthritis,
            CC   for
            CC   diagnostic, prophylactic and therapeutic uses FH   Key
                 Location/Qualifiers
            FT   source          1. .758
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
             |||||  |||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
             ||||||||||:||    ||||||  |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
             |||||  | ||||||||||:||||||:|||||||||||||||  |: :||:|||||||||
Db         117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
             |||||||||||||  |  |||||  ||||||||||||| ||||||||||||||  |:|||||
Db         177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
             |||||  |||||||   :|| | |  || |||||||||||||||  ||||||||||||||||
Db         237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
             ||||||    |||||||||||||||||||| ||||||||| ||| ||||||  |||||||||
Db         297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
             ||||  ||||||  ||||  |||||||||||||||||||||||||  |||||| ||||  ||||
Db         353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
```

```
                |||||||||||| ||||| ||||||||| |||||  | || ||| |||||||||||||||:|
Db        413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
                |||||||   |||||||||||| |||||| |||||||||  |  ||||  ||||:|||||||
Db        473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
                ||||||  |||||  |||||||    |||| | ||||||||||||||||||||||||||||||
Db        530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
                 |||||  ||||||  ||||  |||  ||||||||||||||||||||||  | | |  |||||
Db        590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
                |||||  ||||  ||||  |||||||| ||||||||||||||: ||||||||
Db        710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 264
CS248648
LOCUS       CS248648                758 bp    DNA     linear   PAT 09-JAN-2006
DEFINITION  Sequence 16 from Patent EP1612270.
ACCESSION   CS248648
VERSION     CS248648.1  GI:84662041
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Parahnos-Baccala,G.
  TITLE     Retroviral nucleic material and nucleotide fragments, in particular
            associated with multiple sclerosis and/or rheumatoid arthritis, for
            diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: EP 1612270-A 16 04-JAN-2006;
            BIOMERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
                |||||  |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
                ||||||||||:||    ||||||  |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
                |||||  | ||||||||||:||||||:|||||||||||||| |: :||:|||||||||||
Db         117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
                |||||||||||||  |  |||||  |||||||||||||  |||||||||||||||| |:|||||
Db         177 GATTTTACCCACATGGCCAAATGTCAGGGATTTCAGCATCTACTAGTCTGGGCAGATACT 236

Qy       4786 TTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAG 4845
                |||||  |||||||  :|| | |  || ||||||||||||||| ||||||||||||||||||
Db         237 TTCACTGGTTGGGTGGAGTCTTCTCCTTGTAGGACAGAAAAGACCCAAGAGGTAATAAAG 296

Qy       4846 GCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAT 4905
                ||||||    ||||||||||||||||||| |||||||||| ||| |||||| ||||||||
Db         297 GCACTA----ATGAAATAATTCCCAGATTTGGACTTCCCCCAGGATTACAGGGTGACAAT 352

Qy       4906 AGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATAT 4965
                ||||  |||||| ||||  |||||||||||||||||||||||||| |||||| |||| ||||
Db         353 GGCCCCGCTTTCAAGGCTGCAGTAACCCAGGGAGTATCCCAGGTGTTAGGCATACAATAT 412

Qy       4966 CACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYA 5025
                ||||||||||  |||||  ||||||||  |||||  | || ||| |||||||||||||:|
Db         413 CACTTACACTGTGCCTGGAGGCCACAATCCTCCAGAAAAGTCAAGAAAATGAATGAAACA 472

Qy       5026 CTCAAAGGACATCTAAAAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTTG 5085
                |||||||   |||||||||||| |||||| |||||||||  |  ||||  ||||:|||||||
Db         473 CTCAAAG---ATCTAAAAAAGCTAACCCAAGAAACCCACATTGCATGACCTGTTCTGTTG 529

Qy       5086 CCTATAGCCTTAAAAAGAATCTGCAACT-TTCCCCAAAAAGCAGGACTTAGCCCATACGA 5144
                ||||||  |||||  |||||||    |||| | ||||||||||||||||||||||||||||||
Db         530 CCTATAACCTTACTAAGAATCCATAACTATCCCCCAAAAAGCAGGACTTAGCCCATACGA 589

Qy       5145 AATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA-AGACAGCCAA 5203
                 |||||  ||||||  ||||  |||  ||||||||||||||||||||||  | | |  |||||
Db         590 GATGCTATATGGATGGCCTTTCCTAACCAATGACCTTGTGCTTGACTGAGAAATGGCCAA 649

Qy       5204 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTACAAGG 5263
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db         650 CTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATCACAGGG 709

Qy       5264 AACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
                |||||  ||||  ||||  |||||||| ||||||||||||||: ||||||||
Db         710 AACCTGTCCCCGAGAGGAGGGAAAGGAACTATTCCACCCTGGTGACATG 758


RESULT 265
AR344392
LOCUS       AR344392                758 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 112 from patent US 6582703.
ACCESSION   AR344392
VERSION     AR344392.1  GI:33740333
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 112 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="unknown"
                     /mol_type="mRNA"

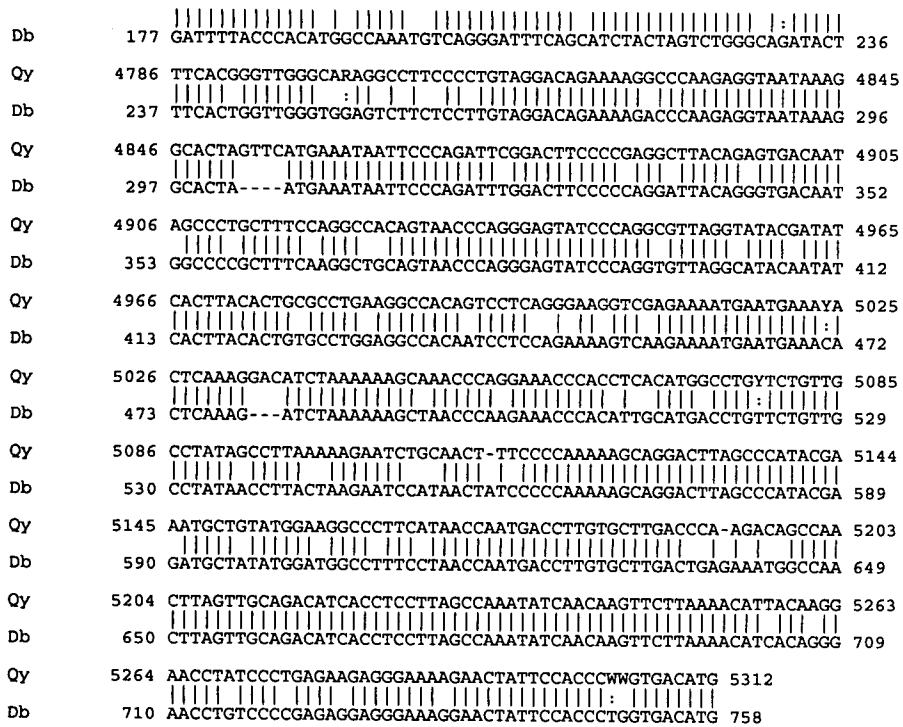
  Query Match             7.5%;  Score 564.8;  DB 1;  Length 758;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  666;  Conservative   16;  Mismatches   73;  Indels   14;  Gaps    6;

Qy       4547 GGCATCGATAGCACCCATCARATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 4606
                |||||  |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 GGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCAAA 60

Qy       4607 ACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCARARAAATAATCCCCTGCCTYAT 4666
                ||||||||||:||    ||||||  |||||:  |||||:|:|||||||||||||||:||
Db         61 ACTATCAAGCAGAT----AGGGCCCGTGAAGCATGCCAAAGAAATAATCCCCTGCCTTAT 116

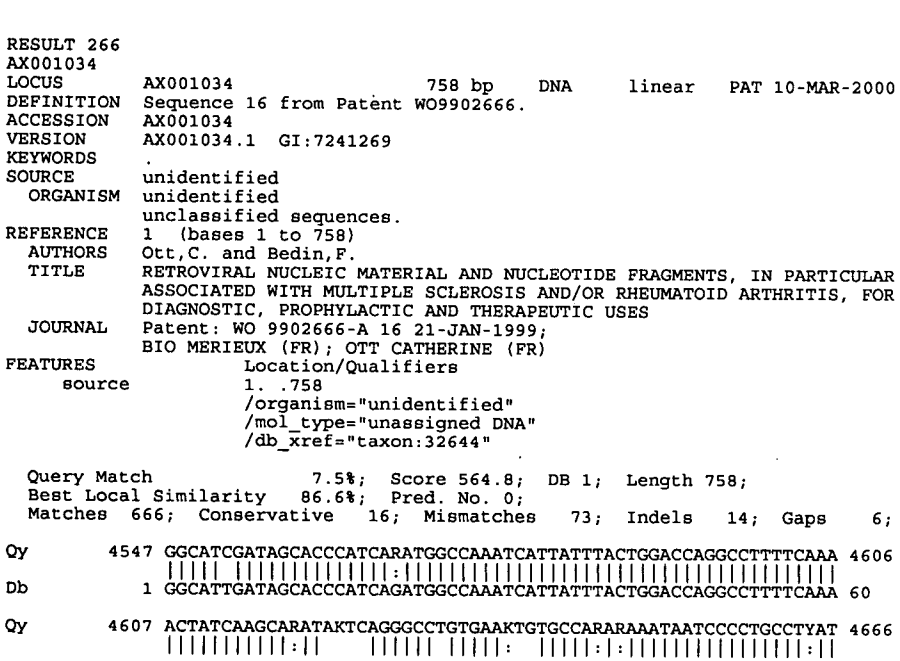
Qy       4667 CGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARACTGGCAACT- 4725
                |||||  | ||||||||||:||||||:|||||||||||||| |: :||:|||||||||||
Db         117 CGCCATGTTCCTTCAGGAGAACAAAGAACAGGCCATTACCCAGGGGAAGACTGGCAACTA 176

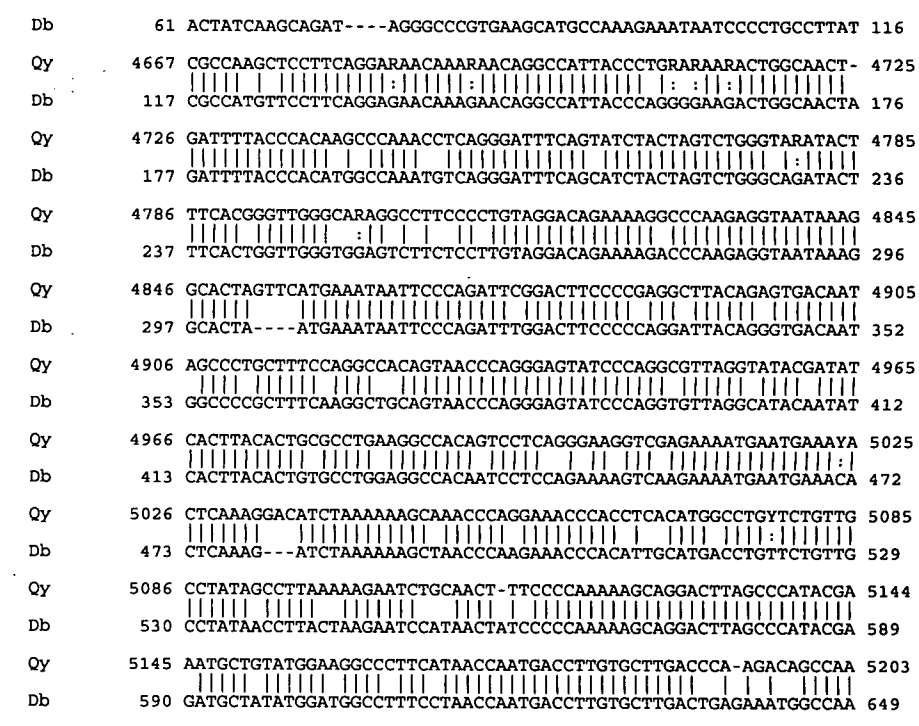
Qy       4726 GATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTARATACT 4785
```



RESULT 266
AX001034
LOCUS AX001034 758 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 16 from Patent WO9902666.
ACCESSION AX001034
VERSION AX001034.1 GI:7241269
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 758)
AUTHORS Ott,C. and Bedin,F.
TITLE RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
JOURNAL Patent: WO 9902666-A 16 21-JAN-1999;
BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES Location/Qualifiers
source 1. .758
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 7.5%; Score 564.8; DB 1; Length 758;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 666; Conservative 16; Mismatches 73; Indels 14; Gaps 6;





RESULT 267
AB115218
LOCUS AB115218 638 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:WP8-1.
ACCESSION AB115218
VERSION AB115218.1 GI:41016648
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 638)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers

```
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP8-1"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             7.5%;  Score 562.4;  DB 1;  Length 638;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  577;  Conservative   60;  Mismatches    1;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||     |||||||:||||:|||||||
Db             1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||:|
Db            57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db           117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||:|||||:|||||:||||||||||||||||||||||
Db           177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:||||||||:|:||||||:::||||||||||||||||||||||||||
Db           236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              |||||||:||||||||||||:||||||||:|||||||||||||||:||:|||:
Db           289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||||:|||:||||||:||||||
Db           342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:||||||||||||||||||||||||:|:||||||
Db           401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||||||||||||::||||||||||||||||||||||||||||||||
Db           461 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||||||||||||:|||||::||||||||||||||||||||||
Db           521 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:||||||||||:||||||||||||||||||:|||||
Db           581 CTTGTGGCTGTCAGACAACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 268
AB115187
LOCUS       AB115187                655 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP1-2.
ACCESSION   AB115187
VERSION     AB115187.1  GI:41016617
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 655)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP1-2"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
                     /note="HERV-W pol"
     repeat_region   1. .655
                     /note="HERV-W pol"

  Query Match             7.5%;  Score 561.4;  DB 1;  Length 655;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  573;  Conservative   57;  Mismatches   23;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              ||||||| |:| |||||:| ||||||||||||||| ||||||||||:||||:|||||||
Db             1 TCCAAACAAAATGCTCAGCACTGCTCACAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db            61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db           121 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||:|||||:|||||:||||||||||||||||||||||
Db           181 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db           240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:||||| | :|||||||||||||:||:|||:
Db           300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :||| | |||||| |||:||:||||||:|||||||||||||||:|||:||||||:||||||
Db           358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||::||||||:||||||||:|||||||||||||||||||||:|:||||||
Db           417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||||||| |||||||||||||
Db           477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| ||||||||||||||||||||:|||||::|||||| |||||||||||||||||
Db           537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
              ||||||||||||||||||:||||||||||:|||||||||||||||||| |:|| |
Db           597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGGAAGGGC 653


RESULT 269
AB115198
```

```
LOCUS       AB115198                638 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-3.
ACCESSION   AB115198
VERSION     AB115198.1  GI:41016628
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP3-3"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             7.5%;  Score 560.8;  DB 1;  Length 638;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  576;  Conservative   60;  Mismatches    2;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||||||||||||||||    ||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db        117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||||:|||||:|||||:||||||||||||||||||||||
Db        177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::| ||||||||||||||||
Db        236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAGGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                    ||||||||:||||||||||||:||||||||:||||||||||||:||:|||:
Db        289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||||:|||:|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:||||||||:||||||||||||||||||||:|:||||||
Db        401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:||||||||||||||||||::||||||||||||||||||||||||||||||||||
Db        461 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||||||||||||||:|||||::||||||||||||||||||||
Db        521 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||:||||||||||:||||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 270
BV607335
LOCUS       BV607335                743 bp    DNA     linear   STS 15-APR-2005
DEFINITION  S217P6762RE12.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV607335
VERSION     BV607335.1  GI:62590520
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 743
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
```

```
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>743

  Query Match           7.5%;  Score 560.6;  DB 1;  Length 743;
  Best Local Similarity   87.8%;  Pred. No. 0;
  Matches  650;  Conservative   19;  Mismatches   59;  Indels   12;  Gaps    6;

Qy       2344 AAAGCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCT 2403
             |||||||||| |||| |||||||||||||| |||||||||| ||  |    ||||
Db         15 AAAGCAGGGGCTATTACACACCTGAACATAGGAAAAGGAACACCCCGTTTGTTGTTCCCT 74

Qy       2404 GCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAA 2463
             |||||||||||||||||||||||||||||| |||||||||| |||||| ||| |||
Db         75 GCTTGAGGAAGGAATTAATCCTGAAGTCTGGGTAACAGAAGGACATTATGGATGAG-CAA 133

Qy       2464 AGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGC 2523
             ||||||||| |||||||||||| |||||||||||||| | ||||||||||||||||||||
Db        134 AGAATGCCCATCCTGTTCAAGTCAAACTAAAGGATTCCGCCTCCTTTCCCTACCAAAGGC 193

Qy       2524 AGTACCCCCTCAGA-CCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAA 2582
             |||| ||||| ||| |||||||||||||||||| ||||||||||||||||||||| |||||
Db        194 AGTATCCCCTTAGACCCCAAGGCCCAACAAGGACTCCAAAAGATTGTTAAGGACCTAAAA 253

Qy       2583 GCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGA 2642
             |||||||| |||||||||||||| | ||||||| || | |         |||   |
Db        254 GCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC-------AATTTTAGG 306

Qy       2643 GGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAG 2702
               | |||||||||| |||||||||| |||||||||||||| ||||||||||||||| |||
Db        307 AGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGACCTCAGGATTATCAAT-GAG 364

Qy       2703 GCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCA 2762
             ||| |||||| | |||||||||||||||||| |||||| | | ||::|||||||||||||
Db        365 GCCATTGTCCCTCTATACCCAGCTGTACCTAACCCTTACATTCTGCTTTCCCAAATACCA 424

Qy       2763 GAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTA 2822
             |||||||||||||||||||:|||||||||||:|||||||||| ||||||||||||||||
Db        425 GAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTA 484

Qy       2823 CATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTC 2882
             |||||||||||||||||||||||||||||||||||||||||| |||||||||| :|||||||||||
Db        485 CATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTCAAACCCCACATCTCAACTC 544

Qy       2883 ACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTA 2942
             |||||||||:|||||||  |||||||||||||||:|||||||||||||||||||| ||||
Db        545 ACCTGGACTGTTTTACCGAAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGTATTA 604

Qy       2943 GCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGAT 3002
             |||||| ||| |||:||:|::|||||||||||||||||||||| :||| |||||||||
Db        605 GCCCAAAACTCGAGCCAATTCTCATACCTGGACACTCTTGTCCTTTGGTACGTGGATGAT 664

Qy       3003 TTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAAT 3062
             ||||||||:||:|||::|||||||||||||||||||||||||||||| :|||||:||
Db        665 TTACTTTTAGCTGCCTGTTCAGAAACCTTGTGCCATCAAGCCACCCAACCGCTCTTAAA- 723

Qy       3063 TTCCTCGCYACCTGTGGCTA 3082
             |||||| |:|||| ||||||
Db        724 TTCCTCACTACCTATGGCTA 743


RESULT 271
BD146807
LOCUS       BD146807                 745 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD146807
VERSION     BD146807.1  GI:27852565
KEYWORDS    JP 2002191363-A/1650.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 1650 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/1650
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   source          1. .745
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match           7.5%;  Score 560.3;  DB 1;  Length 745;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  635;  Conservative   62;  Mismatches   47;  Indels   11;  Gaps    8;

Qy       2842 TGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCC 2901
             |||||||||||||||| ||||||||||:  |||||| ||||||||||:|||||||||
Db          1 TGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAATTCACCTGGACTGTTTTACCCC 60

Qy       2902 AAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCART 2961
             |||||||||||| ||:|||||||||||||||||||||||||||||||||||:||:|
Db         61 AAGGGTTCAGGGACAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGT 120

Qy       2962 YMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTT 3021
             ::||| |||||||||||||||||||:||| |||||||||||||||:|:|||: ||
Db        121 TCTCACACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCTGCCCTTT 180

Qy       3022 CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCT 3081
             ||||||||||||||||||||||||||||||:||||:|||||||| ||:|||||||||
Db        181 CAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTTGCCACCTGTGGCT 240

Qy       3082 ACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAA 3141
             || :|||||||||||:|:| |||||:|||||||||||||| ||||||||||:||||
Db        241 ACCAGGTTTCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAA 300

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
             :|||||| |||||||||:|||||||||||||||:::| |||||| | ||||||||||
Db        301 AATTATCCTAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATC 360

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAA 3261
             ||||||:| |||||||||||||||||:|::|||||| ||:|||:|||:|||||| ||
Db        361 CTCATCCC-AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATAACAGGCTTCTGCCAAA 419

Qy       3262 WATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACT 3321
             :||||||| |||||||:||| :||||||||||:|||||:|||||:|||||||||||||
Db        420 TATGGATT-CCCAGGTATGGTGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACT 478

Qy       3322 CAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCC 3381
             ||||||| |||||||||||||:|||| |||| |:|||||: :||| |||||||||||
Db        479 CAGAAAG-CAATACCCATTTAGTAAGGTGGAAACCTGAAGCGGAAGCAGCTTTCCAGGCC 537
```

```
Qy       3382 CCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTT 3441
Db        538 CTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT-- 593

Qy       3442 STTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGG 3501
Db        594 -CGTATATATCACAG-AAAAAACAGGGATAGCTCTAGGAGTCCTTACACATGTCTGAGAG 651

Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT-GGCAAAGGGTTG 3560
Db        652 ACGAGCTTGC-ACCTGTGGCATACCTGAGTAAGAAAACTGATGTANTGGGCNAAAGGNTG 710

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
Db        711 GCCTCATTGGTTATGGGTAATGGCAGCAATAGCAG 745


RESULT 272
AX866745
LOCUS       AX866745                745 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 1650 from Patent EP1074617.
ACCESSION   AX866745
VERSION     AX866745.1  GI:40021094
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 1650 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             7.5%;  Score 560.3;  DB 1;  Length 745;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  635;  Conservative   62;  Mismatches   47;  Indels   11;  Gaps    8;

Qy       2842 TGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCC 2901
Db          1 TGTTTGCCTTTGAAGATCCTTCAAACCCAATGTCTCAATTCACCTGGACTGTTTTACCCC 60

Qy       2902 AAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCART 2961
Db         61 AAGGGTTCAGGGACAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGT 120

Qy       2962 YMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTT 3021
Db        121 TCTCACACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGCTGCCCTTT 180

Qy       3022 CAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCT 3081
Db        181 CAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTAAATTTCCTTGCCACCTGTGGCT 240

Qy       3082 ACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAA 3141
Db        241 ACCAGGTTTCCAAACCAAACGCTCAGCTCTGCTCACAGCAGGCTAAATACTTAGGGCTAA 300

Qy       3142 RATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATC 3201
Db        301 AATTATCCTAAGGCACCAGGGCCCTCAGTGAGGAATGTACCCAGCCCACACTGGCTTATC 360

Qy       3202 CTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAA 3261
Db        361 CTCATCCC-AAACCCTAAAGCAACTAAGAGGATTCCTTAGCATAACAGGCTTCTGCCAAA 419

Qy       3262 WATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACT 3321
Db        420 TATGGATT-CCCAGGTATGGTGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACT 478

Qy       3322 CAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCC 3381
Db        479 CAGAAAG-CAATACCCATTTAGTAAGGTGGAAACCTGAAGCGGAAGCAGCTTTCCAGGCC 537

Qy       3382 CCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTT 3441
Db        538 CTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT-- 593

Qy       3442 STTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGG 3501
Db        594 -CGTATATATCACAG-AAAAAACAGGGATAGCTCTAGGAGTCCTTACACATGTCTGAGAG 651

Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGT-GGCAAAGGGTTG 3560
Db        652 ACGAGCTTGC-ACCTGTGGCATACCTGAGTAAGAAAACTGATGTANTGGGCNAAAGGNTG 710

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAG 3595
Db        711 GCCTCATTGGTTATGGGTAATGGCAGCAATAGCAG 745


RESULT 273
BV600201
LOCUS       BV600201                757 bp    DNA     linear   STS 14-APR-2005
DEFINITION  S217P60751RG2.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV600201
VERSION     BV600201.1  GI:62539115
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 757
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
```

```
          score >= 25), and the read must have at least 200 bp SNQS(30,25)
          bases. Reads not uniquely
          placed in the genome and read pairs whose two ends were not
          consistently placed were
          discarded. After above filtering, NQS(30,25) standard was applied
          to all pairs of
          overlapping reads to call NQS bases and SNPs. Alignments (between
          two reads) with less
          than 100 NQS bases or with SNP rate > 0.01 were discarded. To
          exclude alignment between two
          copies of a single read, comparisons between two reads that share
          95% of their genome
          alignments (>=95% bases of read A and >=95% bases of read B were
          placed at the same locus
          of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>757

  Query Match             7.5%;  Score 558.3;  DB 1;  Length 757;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  626;  Conservative   61;  Mismatches   69;  Indels    7;  Gaps    6;

Qy       3158 CARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCT 3217
              ||:|||||||||||||| |:::|||||||||| ||||||| |||| |||:||||||| |
Db             1 CAGGGCCCTCAGTGAGG-ACATATCCAGCCTATCCTGGCTTTTCCTTATCCCAAAACCGT 59

Qy       3218 AAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGT 3277
              ||||||||||||:  :|||||||||:|||:|||:|||||| |||:||||||| ||| |||
Db            60 AAAGCAACTAAGAACGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT-CCCTGGT 118

Qy       3278 WTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCC 3337
              :  ||:|||||||||| :|||||:|||||::|||||||||||||||||||||| ||||||||
Db           119 ACAGCAAAATAGCCAGATCATTATATACACTAATTAAGGAAACTCAGAAAGGCAATACCC 178

Qy       3338 ATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTA 3397
              |||||:|||||||||:|:||||||:::|||||| ||||||||||||   ||  | | |||
Db           179 ATTTAGTAAGATGGACACCTGAAGCAGAAGTGACTTTCCAGGCCCTAAAGAAGGACCT-- 236

Qy       3398 AACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGA 3457
              ||||||||:|||||||||||:||||||||: |||||||||||||:||:|||::||||||
Db           237 AACCCAAGCCCCAGTGTTAAGCTTGCCAACGTGGCAAGACTTTTCTTTATATGTCACAG- 295

Qy       3458 AAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYR 3517
              |||||||||:||:||||||:|||||||||||   |||:|||:| |||: ||||||||||::
Db           296 AAAAAACAGGAATAGCTCTAGGAGTCCTTATGCAGGTCCGACGGACCAGCTTGCAACCTG 355

Qy       3518 TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGG 3577
              ||||::|||| |: |||||||:|||||||||||||||||||:|:|||| |||||:|||
Db           356 TGGCATACCTAAGCAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATCGTTTATGGG 415

Qy       3578 TAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATC 3637
              ||||||  |||||||||||  :|||||||||||||||||||||||| ||||||||:|||||||
Db           416 TAGTGGCAGCAGTAGCAGTATTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGATC 475

Qy       3638 TTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGC 3697
              |||||||||| |||||||| ||:|||||::|||||||||||||||||||||||||||||||
Db           476 TTACTGTGTGAACATCTCAAGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGGA 535

Qy       3698 TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRAC 3757
              ||||||||||: ||| |||||||:|||| ||||||||||||||:||||||||||||:||
Db           536 TGTCAGACAACCATTTGCTTAAATATCAGCATCTATTACTTGAAGGGCCAGTGCTGCGAC 595

Qy       3758 TGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATAR 3817
              ||||| ||||| |||||||||| ||||:| |||||||||| |  |    |||| | |||:
Db           596 TGTGCCCTTGTACAACTCTTAATCCAGCCACATTTCTTCCTGGTACATGAGAAGAAATAG 655

Qy       3818 AAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARG 3877
              ||:|||||||||    |:||| | |||||||||| ||     ||| ||||||:||||:|
Db           656 AACATAACTGTCCCACAGTAACTGCTCAAACCTACACCCACTCAGGTGACCTTCTAGAGG 715

Qy       3878 TTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA 3920
              |||: ||||||||| :|||| || |||||||| |||||||||
Db           716 -TCCCCTGACTGATCATGACCCTC-ACTTGTATGCTGATGGAA 756


RESULT 274
AB115233
LOCUS       AB115233                656 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-9.
ACCESSION   AB115233
VERSION     AB115233.1  GI:41016663
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP10-9"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
                     /note="HERV-W pol"
     repeat_region   1. .656
                     /note="HERV-W pol"

  Query Match             7.4%;  Score 554.2;  DB 1;  Length 656;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  567;  Conservative   60;  Mismatches   28;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||||| |||||||||||||||||||| |:||||:|||||||
Db             1 TCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAAGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||| |:|||||||||||||:::||||||| |||||||||||||||||||:|
Db            61 AAAGGCACTAGGGCCCTCAGTGAGGAATGTATCCAGCATATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db           121 AAAACCCTAAAGCAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:|||||| ||||:|||||:|||||:|||||||||||||||||||||||
Db           181 -CCCAGGTATGGCAAAATAGTCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||:|:||||||:::||| ||||||||||||||   |||
Db           240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCC--TAAGG 297

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | | |   ||||||||:||||||||||||:||||||||:||||||||||||||:||:|||:
Db           298 AAGGCCCTAACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAC-AGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
```

```
            :|||||||||||||  ||:||:||||||:||||||||||||||||:|| :|||||: ||||
Db        358 GTCACAGAAAAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACCAGCT 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::|| |||:|||| |||:||||||| |||||||||||||:|:|||||
Db        418 TGCAACCTGTGGCATACTTGAGTAAGAAAATTGATGTACTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||:|||||||||   ||||||||||::|||||||||||||||||||| |||||||||
Db        478 GTTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |:|||||||||||||||| ||||||||||:|||||::|||||||||||||||||| ||
Db        538 GAAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGAAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||| |||||||||||||||:|||| |||||||:|||||||||||||||||:|||||
Db        598 ACTGGTGGCTGTCAGACAACTGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 656
```

```
RESULT 275
BV595108
LOCUS       BV595108                 746 bp    DNA     linear   STS 14-APR-2005
DEFINITION  S215P61466RH4.T0 Clara Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV595108
VERSION     BV595108.1  GI:62534022
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 746
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .746
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Clara"
     STS             <1. .>746

  Query Match             7.4%;  Score 553.7;  DB 1;  Length 746;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  640;  Conservative   31;  Mismatches   54;  Indels   25;  Gaps    6;

Qy       2463 AAGAATGCCCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGG 2522
             | |||||||| ||||||||||||  |  || |||||||  | |||||||||||||||||||
Db         21 AGGAATGCCCATCCTGTTCAAGTCTATATAGAGGATTCTGCCTCCTTTCCCTACCAAAGG 80

Qy       2523 CAGTACCCCCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAA 2582
             ||||||   || |||||| |||||||||||||| | ||||||||||||||||||| |||||
Db         81 CAGTACTTTCTTAGACCCGAGGCCCAACAAGGACT-CAAAAGATTGTTAAGGACCTAAAA 139

Qy       2583 GCCCAAGGCTTAGTAAAACCATGCATAACTCCCTGCAGTAATTCCGTAGTGGATTGAGGA 2642
             |||||||| |||||||||||||||  |  ||||||| || | |        |||  |
Db        140 GCCCAAGGCCTAGTAAAACCATGCAATAGCCCCTGCAATACTCC-------AATTTTAGG 192

Qy       2643 GGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTATCAATGGAG 2702
              | |||||||||||  |||||||||| || |||| |||||||||||||||||||||| |||
Db        193 AGTACAGAAACCCAACGGACAGTGGA-GGCTAGTACAAGATCTCAGGATTATCAAT-GAG 250

Qy       2703 GCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCCAAATACCA 2762
             |||||||| | |||||||||||||||||| ||||||||||||::|||||||||||||||
Db        251 GCCGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACTGTGCTTTCCCAAATACCA 310

Qy       2763 GAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCATCCCTGTA 2822
             |||||| |||||||||||||:||||||||||:|||||||||| ||||||||||||||||
Db        311 GAGGAAACAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCATCCCTGTA 370

Qy       2823 CATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCATCTCAACTC 2882
             |              |||||||||||||||||||| ||||||||||:||||||||||||
Db        371 C--------------TTCTTGTTTGCCTTTGAAGATCCTTCAAACCCAACATCTCAACTC 416

Qy       2883 ACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTA 2942
             |||||||||:|||||||||||||||||||||||:|||||||||||||||||| ||||
Db        417 ACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGAATTA 476

Qy       2943 GCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGAT 3002
             |||||||||||||||:||:|::|||||||||||||||||||||||||:|||:||||||||
Db        477 GCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCGGTATGTGGATGAT 536

Qy       3003 TTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAAT 3062
             |||||| |:||: ||::|||||||||||||||||||||||||||||||||:||||| |||
Db        537 TTACTTCTAGCCACCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCGCTCTTTAAT 596

Qy       3063 TTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAG 3122
             ||||||||:|||||||||||||:||||||||||:|:|:|||||:|||||||||| |||
Db        597 TTCCTCGCCACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAACAG 656

Qy       3123 GTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATC 3182
             |||||||||||  ||:||||:| |||||||||  |||||:|||||||||||||| :::|||
Db        657 GTTAAATACTTCGGGCTAAAA-TATCCAAAGACACCAGGGCCCTCAGTGAGGACTGTATC 715

Qy       3183 CAGCCTATACTGGCTTATCCTCATCYCAAA 3212
             ||||| |||||||  ||||||||||:||||
```

```
Db        716 CAGCCCATACTGGGCTATCCTCATCCCAAA 745


RESULT 276
AB056502
LOCUS       AB056502                654 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWPX-1.
ACCESSION   AB056502
VERSION     AB056502.1  GI:15341230
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWPX-1"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .654
                     /gene="pol"
     misc_feature    <1. .>654
                     /gene="pol"

  Query Match             7.4%;  Score 552;  DB 1;  Length 654;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  569;  Conservative   55;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||||| ||||||||:||||:|||||||
Db          1 TCCAAACGAAAGGCTCAGCTCTGCTCACAGCAGGTTAAGTACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||||| ||||||||||||||:::|||||||||| ||||||||||||||||||:|
Db         61 AAAGGCACCACGGCCCTCAGTGAGGAATGTATCCAGCCTAAACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||||||||||:| :|||||||||:|||:|||:||||||| ||: ||||||
Db        121 AAAACCCTAAAGCAACTAAGAGCATTCCTTGGCATAACAGGCTTCTGCCAAATGTGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:  ||:||||||||||:|||||:|||||:|||||||||||||||||||||||
Db        181 -CCCAGGTACAGCAAAATAGCCAGGCCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:||||||||:|:||||||:::|||  ||||||||  |||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCAGCTTTCCATACCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:||||||||:||||||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||||| |||| ||:||:|||||:||||||||||||||||:|||:|||||:||||||
Db        358 GTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::|||||| |||||||| |||||||| ||||||||||:|:||||||
Db        417 GCAACCTGTGGCATACCTGAATAAGGAAAATGATGTAGTAGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||||||||||||| |||||
Db        477 TTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATGCAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||| ||||||||||||||||||||||:|||||::|||||| ||||||||||||||
Db        537 AAGAGCTCTTACTGTGTGGACATCTCATGATGTGAATGGCATACCCACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||  ||||||||||:|||||| ||||||||||:|||||
Db        597 CTTGTGGCTGTCAGACAACAATTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 654


RESULT 277
AB115291
LOCUS       AB115291                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP10-6.
ACCESSION   AB115291
VERSION     AB115291.1  GI:41052519
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP10-6"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.4%;  Score 551.2;  DB 1;  Length 654;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  567;  Conservative   57;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              ||||||| |:| |||||:| ||||||||||||||||||| |||||||:||||:|||||||
Db          1 TCCAAACAAAATGCTCAGCACTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||| |:||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||||||||||:|::|||||||||:|||:| |:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: || :|||||| ||| :|||| :|||||:||||||||||||  |||||
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239
```

```
Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGC 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654
```

RESULT 278
AB115193

| | |
|---|---|
| LOCUS | AB115193 654 bp RNA linear PRI 01-FEB-2006 |
| DEFINITION | Homo sapiens RNA transcript, HERV-W pol, isolate:WP2-5. |
| ACCESSION | AB115193 |
| VERSION | AB115193.1 GI:41016623 |
| KEYWORDS | . |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 |
| AUTHORS | Yi,J.M., Kim,H.M. and Kim,H.S. |
| TITLE | Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells |
| JOURNAL | J. Gen. Virol. 85 (PT 5), 1203-1210 (2004) |
| PUBMED | 15105536 |
| REFERENCE | 2 (bases 1 to 654) |
| AUTHORS | Yi,J. and Kim,H. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962) |

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .654<br>/organism="Homo sapiens"<br>/mol_type="other RNA"<br>/isolate="WP2-5"<br>/db_xref="taxon:9606"<br>/tissue_type="prostate"<br>/note="HERV-W pol" |
| repeat_region | 1. .654<br>/note="HERV-W pol" |

Query Match 7.3%; Score 550.4; DB 1; Length 654;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 565; Conservative 59; Mismatches 30; Indels 4; Gaps 3;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db          61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db         121 AAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db         181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db         240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db         300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db         358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db         417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db         477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db         537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db         597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654
```

RESULT 279
AB115209

| | |
|---|---|
| LOCUS | AB115209 654 bp RNA linear PRI 01-FEB-2006 |
| DEFINITION | Homo sapiens RNA transcript, HERV-W pol, isolate:WP6-3. |
| ACCESSION | AB115209 |
| VERSION | AB115209.1 GI:41016639 |
| KEYWORDS | . |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 |
| AUTHORS | Yi,J.M., Kim,H.M. and Kim,H.S. |
| TITLE | Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells |
| JOURNAL | J. Gen. Virol. 85 (PT 5), 1203-1210 (2004) |
| PUBMED | 15105536 |
| REFERENCE | 2 (bases 1 to 654) |
| AUTHORS | Yi,J. and Kim,H. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962) |

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .654<br>/organism="Homo sapiens"<br>/mol_type="other RNA"<br>/isolate="WP6-3"<br>/db_xref="taxon:9606"<br>/tissue_type="liver" |

```
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 550.4;  DB 1;  Length 654;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 280
AB115228
LOCUS       AB115228                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-1.
ACCESSION   AB115228
VERSION     AB115228.1  GI:41016658
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP10-1"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 550.4;  DB 1;  Length 654;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 281
AB115281
LOCUS       AB115281                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP8-4.
ACCESSION   AB115281
VERSION     AB115281.1  GI:41052509
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP8-4"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 550.4;  DB 1;  Length 654;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGGTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                ||||||||||||||||||:||||||||||:||||||||||||||||||:|||||
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 282
BV633838/c
LOCUS       BV633838               716 bp    DNA     linear   STS 15-APR-2005
DEFINITION  S216P61314RD11.T0 Masuku Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV633838
VERSION     BV633838.1  GI:62624016
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 716
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
```

```
                /clone_lib="Masuku"
     STS        <1. .>716

  Query Match             7.3%;  Score 549.1;  DB 1;  Length 716;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  637;  Conservative    5;  Mismatches   64;  Indels    5;  Gaps    5;

Qy       1414 GCTGCAGACATTTGCTAACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCT 1473
             |||||||||||||    ||||||  |||||:|||||||  ||| |||||||:|||| ||: ||
Db        716 GCTGCAGACATTGCTAAACTTGCATGCTAGAAGGACTTAGGGAAACTAGGAAG-AAGCCT 658

Qy       1474 AYGAATTACTCAATGATGTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTT 1533
             |: |||    ||||||||||||||| || |||||||| || ||||||||||||||||||||
Db        657 ATAAATATTTCAATGATGTCCACTAT-ACACAGGGAAAAGGAAGAAAATCCTACTGCCTT 599

Qy       1534 TCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGA 1593
             |||||||||||||||||||||||||||||| | | ||||||||||||||||||  |||
Db        598 TCTGGAGAGACTAAGGGAGGCATTGAGGAAGCATACTTCTCTGTCACCTGACTCTATTGA 539

Qy       1594 AGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAA 1652
             |||||||||||||||||  |||||||||||||| ||||||||||||||||||| |||||
Db        538 AGGCCAACTAATCTTAAAGGATAAGTTTATCACTCTGTCAGCTGCAGACATTAGAAAAAA 479

Qy       1653 ACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAAC 1712
             |||||||||||||||| | | || ||| |||||||||| ||||||||||||||||||||
Db        478 ACTTCAAAAGTCTGCCTTCGACCTGGAACAAAACTTAGATACCCTATTGAACTTGGCAAC 419

Qy       1713 YTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAA 1772
             :|||||||||||||||| |||||||||||||||||||||||| ||||||| | || ||| |
Db        418 CTCGGTTTTTTATAATAAAGATCAGGAGGAGCAGGCGGAACGGGACAAATGAGA-TAAGA 360

Qy       1773 AAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAG 1832
             |||||||||| |||||||||| |||| ||||||| |||||||||||||||||| || |
Db        359 AAAAGGCCACCACTTTAGTCATGGCCCTTAGGCAAGCGGACTTTGGAGGCTCTGAAACAC 300

Qy       1833 GGAAAAGCTGGGCAAATTGAATGCCTAAT-AGGGCTTGCTTCCAGTGCGGTCTACAAGGA 1891
             |||||  ||||||||||  |||||||||| |||||||||||||||||||||| |||||||
Db        299 GGAAAGGCTGGGCAAATGAAATGCCTAATAAGGGCTTGCTTCCAGTGCGGTCGACAAGGA 240

Qy       1892 CACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTT 1951
             |||||||||||||||||||| ||||| |||||| |||| ||||||||||||||||| |
Db        239 CACTTTAAAAAAGATTGTCCAAATAGAAATAAGCCACCCCCTCGTCCATGCCCCTTATGT 180

Qy       1952 CAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCC 2011
             |||||||||||||||||||||||||| ||||||| |  |||||| | ||||||||||||||
Db        179 CAAGGGAATCACTGGAAGGCCCACTGGCCCAGGGAATGAAGGTCCTCTGAGTCAGAAGCC 120

Qy       2012 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCC 2071
             |||||||||||||||||||||||||||||||||||  ||||||| |||| ||||||||||
Db        119 ACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCCAGGGCAAGTGCCAGCCCATGCC 60

Qy       2072 ATCACCCTCACAGAGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT 2122
             |||||||||||||||| |||||||||||||||||||||||||||| | |
Db          59 ATCACCCTCACAGAGCCCCGGGTATGCTTGACCATTGAGGGCCAGGAGGTT 9


RESULT 283
AB115249
LOCUS       AB115249                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP1-3.
ACCESSION   AB115249
VERSION     AB115249.1  GI:41052477
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
    PUBMED  15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP1-3"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 548.8;  DB 1;  Length 654;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  564;  Conservative   59;  Mismatches   31;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             ||||| |:|:|:| |||:|||||||| |||||||| ||||||||:||||:|||||||
Db          1 TCCAACCCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||| |||||||||||||:|::|||||||||:|||:|||:|||||  ||||:||||||||
Db        121 AAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: || :|||||| ||| :|||| :|||||:|||||||||||||  |||||
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::||| ||||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  |||||||||:|||||||||||||:|||||  | :|||||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :||| | |||||| |||:||:|||||||:|||||||||||||||||:|||:|||||:|||||
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::|||||| ||||||||:|||||||||||||||||||:|:|||||||
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  |||||||||||::|||||||||||||||||| ||||||||||||
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||| |||||||||||||||||||||:|||||::|||||| ||||||||||||||||
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||:|||||||||||:|||||||||||||||||||:|||||
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 284
AB115186
LOCUS       AB115186                638 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP1-1.
ACCESSION   AB115186
```

```
VERSION     AB115186.1  GI:41016616
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP1-1"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 548.4;  DB 1;  Length 638;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  569;  Conservative   59;  Mismatches   10;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGACCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        461 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        521 AAGAGATCTTACTGTGTGAACCTCTCATGATGTGAACCGCATACTCACTGCTAAAGAAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        581 CTGGTGGCTGTCAGACAACTGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 285
AB115197
LOCUS       AB115197                640 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-2.
ACCESSION   AB115197
VERSION     AB115197.1  GI:41016627
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 640)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP3-2"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .640
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 548.4;  DB 1;  Length 640;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  569;  Conservative   60;  Mismatches   11;  Indels   20;  Gaps    4;

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db          1 TTTCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTAT 56

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db         57 CCAAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATC 116

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db        117 CCAAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGA 176

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db        177 TT-CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAA 235

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db        236 GCCAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC----- 290

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db        291 ---------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCAT 341

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
```

```
Db        342 ATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGC 400

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
              ||||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||
Db        401 TTGCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCAT 460

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
              ||||||:|||||  |||  ||||||||||::|||||||||||||||||||||  ||||||
Db        461 TGTTTATGGGTAATGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAG 520

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:||||||||||||||||||  ||||||||:|||  |::||||||||||||||||||||
Db        521 GGAAGAGATCTTACTGTGTGGATGTCTCATGATGTGGACGGCATACTCACTGCTAAAGGA 580

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||  ||||||||||||||:|||| |||||||:|||||||||||||||||:|||||
Db        581 GACTTACGGCTGTCAGACAACCGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 640


RESULT 286
AB115214
LOCUS       AB115214                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-7.
ACCESSION   AB115214
VERSION     AB115214.1  GI:41016644
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-7"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 547.8;  DB 1;  Length 654;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  563;  Conservative   59;  Mismatches   31;  Indels    4;  Gaps    3;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              | ||||:|:|:| |||:|||||||| ||||||||| ||||||||||:||||:||||||||
Db          2 CAAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCCA 61

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:||
Db         62 AAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCCA 121

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||| |||||||||||||:|::|||||||||:|||:|||:||||| ||||:|||||||
Db        122 AAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT- 180

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||: || :|||||| ||| :|||| :|||||:|||||||||||||||  ||||| |

Db        181 CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGTC 240

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||:|||||||||:|:||||||:::|||  |||||||||||||||  ||  |
Db        241 AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAG 300

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:|||||||||||||:|||||  | :|||||||||||||:||:|||::
Db        301 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATATG 358

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              ||| | |||||| |||:||:||||||:|||||||||||||||:|||:|||||:|||||||
Db        359 TCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTTG 417

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
              |||||::||||::|||||| ||||||||:|||||||||||||||||||:|:|||||||
Db        418 CAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 477

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
              |||:|||||||  ||||||||||::|||||||||||||||||  ||||||||||||:
Db        478 TTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGGA 537

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
              |||||| ||||||||||||||||||||||:|||||::|||||| ||||||||||||||
Db        538 AGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGAC 597

Qy       3691 TTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:||||||||||:|||||||||||||||||||:|||||
Db        598 TTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 287
AB115290
LOCUS       AB115290                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP10-5.
ACCESSION   AB115290
VERSION     AB115290.1  GI:41052518
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP10-5"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 547.6;  DB 1;  Length 654;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  564;  Conservative   58;  Mismatches   32;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              ||||||| |:| |||||:| |||||||||||||||||| |||||||:||||:|||||||
```

```
Db          1 TCCAAACAAAATGCTCAGCACTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||| |:|||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||||||||||||:|::|||||||:|||:|| :||||||||||:||||| |
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGTTTTCTGCCGAATATGGACT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              || | ||:  ||:|||||||||| :|||||:|||||:|||||||||||||||||||||
Db        181 -CCTAAGTACTGCGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||| ||||:|||||| :|||||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|| || ||| |||||||:||:|||||||:|||||||||| ||||:|||:|||||: |||||
Db        358 GTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              |||||::||||::|||||:|||||||:|||||| |||||||||||||:|:||||||
Db        417 GCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||||||| ||||||||||
Db        477 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||| ||||||||||:|||||::|||||||||||||||||||||||
Db        537 AAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||:||||||||||||:|||||||||||||||||||:| |||
Db        597 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 654


RESULT 288
AB115215
LOCUS       AB115215                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-8.
ACCESSION   AB115215
VERSION     AB115215.1  GI:41016645
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-8"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 547.2;  DB 1;  Length 654;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  563;  Conservative   59;  Mismatches   32;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:| |||:|||||||| ||||||||| ||||||||||:||||:|||||||
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||| |||||||||||||||:|::|||||||:|||:|||:||||| ||||:|||||||
Db        121 AAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              ||||||| || :|||||| ||| :|||| :|||||:||||||||||||  |||||
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||:||||| | :|||||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :||| | |||||| |||:||:||||||:|||||||||||||||:|||:|||||:|||||
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::|||||| ||||||||:|||||||||||||||||||:|:||||||
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||||||| ||||||||||
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||||||||||||||:||||||::|||||| ||||||||||||||||
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAATGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||| |||||||| |||||:||||||||||||:|||||||||||||||||||:|||||
Db        597 CTTGGGGCTGTCAAACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 289
AB126104
LOCUS       AB126104                654 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP1-12.
ACCESSION   AB126104
VERSION     AB126104.1  GI:47169704
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 654)
```

```
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP1-12"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 547.2;  DB 1;  Length 654;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  563;  Conservative   59;  Mismatches   32;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||||||||||||| ||||||| :||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAATTACTTAGTGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||| ||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCTGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:  ||:|||||||||| :|||||:|||||:|||||||||||||||||||||
Db        181 -CCCAGGTACAGCGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||| ||||||||:|||| ||||:|  |||||::: || |||||||||| ||  ||  |
Db        240 CAATGCCCATTTAGTAAGGTGGACATATGAAGCAGCAGCGGCTTTCCAGGTCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||| ||||||||:||| ||| :|| ||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCTGTGTTAAGCTTGTCAATGGGCCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:|||||:|||||||||||||||:|||:|||||:| ||||
Db        358 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGCGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||| ||:|||||||:|||||||||||||||||:|:||||||
Db        417 GCAACCCGTGGCATACCCGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||||||||||||::|||||||||||||||||||||||||||||||||
Db        477 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||| ||||||||||||:|||||::|||||||| ||||||||||||||
Db        537 AAGAGATCTTACTGTGGGGACATCTCATGATGTGAACGGCATACTCGCTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             || |||| ||||||||||:||||||||||||:|||| ||||| ||||||||:|||||
Db        597 CTCGTGGTTGTCAGACAACTGTTTACTTAAATATCAGCCTCTACTACTTGAAGGGCCA 654


RESULT 290
AB115254
LOCUS       AB115254                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP2-3.
ACCESSION   AB115254
VERSION     AB115254.1  GI:41052482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP2-3"
                     /db_xref="taxon:9606"
                     /cell_line="PFSK-1"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 545.6;  DB 1;  Length 654;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   33;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:| |||:|| ||||  |||||||| ||||||||||:||||:|||||||
Db          1 TCCAAACCAAAAGTTCAGCTTTGCTTTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||| ||||||||||||||:|::|||||||||:|||:|||:||||| ||||:|||||||
Db        121 AAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: || :|||||| ||| :|||| :|||||:|||||||||||||  |||||
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::||| ||||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:|||||  | :|||||||||||||:||:|||:
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :||| | |||||| |||:||:|||||:|||||||||||||||:|||:|||||:|||||||
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||::|||||| ||||||||:|||||||||||||||||||:|:||||||
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  |||||||||::|||||||||||||||| ||||||||||||||||
Db        477 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||| ||||||||||||||||||||::|||||| |||||||||||||||||||||||
Db        537 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596
```

```
Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 654
```

```
RESULT 291
AB115194
LOCUS       AB115194                655 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP2-8.
ACCESSION   AB115194
VERSION     AB115194.1  GI:41016624
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 655)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP2-8"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
                     /note="HERV-W pol"
     repeat_region   1. .655
                     /note="HERV-W pol"

  Query Match             7.3%;  Score 543.4;  DB 1;  Length 655;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  561;  Conservative   58;  Mismatches   34;  Indels    4;  Gaps    3;
```

```
Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db          3 CAAACCAAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCCA 62

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
Db         63 AAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCCA 122

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db        123 AAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT- 181

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
Db        182 CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGGAACTTGGAAAGTC 241

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
Db        242 AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAG 301

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
Db        302 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATATG 359

Qy       3451 TCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
Db        360 TCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTTG 418
```

```
Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
Db        419 CAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 478

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
Db        479 TTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGGA 538

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
Db        539 AGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGAC 598

Qy       3691 TTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        599 TTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 655
```

```
RESULT 292
AB115210
LOCUS       AB115210                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-1.
ACCESSION   AB115210
VERSION     AB115210.1  GI:41016640
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-1"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 542.4;  DB 1;  Length 654;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  560;  Conservative   59;  Mismatches   35;  Indels    4;  Gaps    3;
```

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGCTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239
```

```
Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654
```

```
RESULT 293
AB056495
LOCUS       AB056495                657 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP11-3.
ACCESSION   AB056495
VERSION     AB056495.1  GI:15341223
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP11-3"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .657
                     /gene="pol"
     misc_feature    <1. .>657
                     /gene="pol"

  Query Match             7.2%;  Score 541.7;  DB 1;  Length 657;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  566;  Conservative   60;  Mismatches   28;  Indels    7;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAACTCTGGTCAGAACAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db          61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCTTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db         121 AAAGCCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db         181 -CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db         240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
Db         300 GGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTA 357

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
Db         358 TATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAG 416

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
Db         417 CCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCA 476

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db         477 TTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACA 536

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db         537 GGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGG 596

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db         597 AGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACC 656

Qy       3747 A 3747
Db         657 A 657
```

```
RESULT 294
AB056498
LOCUS       AB056498                657 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP14-2.
ACCESSION   AB056498
VERSION     AB056498.1  GI:15341226
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .657
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP14-2"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .657
                     /gene="pol"
     misc_feature    <1. .>657
                     /gene="pol"

  Query Match             7.2%;  Score 541.7;  DB 1;  Length 657;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  566;  Conservative   60;  Mismatches   28;  Indels    7;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||  | ||||  |||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||||:::|||||||| ||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::||| |||||:|| :|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||:|||||||  |||:|| ||:|||||:|||||  ||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCTAGGCCAATATATACACTAATTTAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
              | |||  ||||||||:|||||||    ||||:|||||||:|||||||||||| :||:|
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTA 357

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| |||||||||:||:||||||:|||||||||||||||:|||:|||||: ||
Db        358 TATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAG 416

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              | |||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:|||
Db        417 CCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCA 476

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||: |||||||| ||||||||||||::|||||||||||||||||| |||||||
Db        477 TTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACA 536

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             || :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||
Db        537 GGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGG 596

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||:||||||||||:|| ||||||| | ||||||:| ||
Db        597 AGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCCGGCTCTACTGCTTGAAGGACC 656

Qy       3747 A 3747
              |
Db        657 A 657


RESULT 295
AB056500
LOCUS       AB056500                657 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP20-1.
ACCESSION   AB056500
VERSION     AB056500.1  GI:15341228
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP20-1"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .657
                     /gene="pol"
     misc_feature    <1. .>657
                     /gene="pol"

  Query Match             7.2%;  Score 541.7;  DB 1;  Length 657;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  566;  Conservative   60;  Mismatches   28;  Indels    7;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||  | ||||  |||||||||||||:||||:|||| ||
Db          1 TCCAAACCAAAAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTACCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||||:::|||||||| ||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::||| |||||:|| :|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||:|||||||  |||:|| ||:|||||:|||||  ||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAGCTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
              | |||  ||||||||:|||||||    ||||:|||||||:|||||||||||| :||:|
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTA 357

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
             ||::|||||| |||||||||:||:||||||:|||||||||||||||:|||:|||||: ||
Db        358 TATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAG 416

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              | |||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:|||
Db        417 CCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCA 476

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||: |||||||| ||||||||||||::|||||||||||||||||| |||||||
Db        477 TTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACA 536

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             || :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||
Db        537 GGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGG 596
```

```
Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             ||||||||||||||||||||||:||||||||||||:|||||||||| | ||||||:| ||
Db        597 AGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACC 656

Qy       3747 A 3747
             |
Db        657 A 657
```

```
RESULT 296
AB115220
LOCUS       AB115220                640 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP8-8.
ACCESSION   AB115220
VERSION     AB115220.1  GI:41016650
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 640)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP8-8"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
                     /note="HERV-W pol"
     repeat_region   1. .640
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 541.3;  DB 1;  Length 640;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  570;  Conservative   60;  Mismatches    7;  Indels   21;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||||||||||    ||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db        117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||:|||||:|||||:||||||||||||||||||||||
Db        177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||||||||||
Db        236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                      ||||||||:||||||||||||:||||||||:|||||||||||||:||:|||:
Db        289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341
```

```
Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||  ||||||||:||:||||||:||||||||||||||||:|||:|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:||||||||:||||||||||||||||||||||:|:||||||
Db        401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA-GG 3628
              ||||:||||||||||||||||||||::||||||||||||||||||||||||||||| ||
Db        461 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 520

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
              |:|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        521 GAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAA 580

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
              |||||||||||||| |||||:||||||||||:|||||||||||||||  |: | |
Db        581 ACTTGTGGCTGTCAAACAACTGTTTACTTAAATGTCAGGCTCTATTACTTTGAAAGGC 638
```

```
RESULT 297
AB056493
LOCUS       AB056493                657 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP8-1.
ACCESSION   AB056493
VERSION     AB056493.1  GI:15341221
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP8-1"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .657
                     /gene="pol"
     misc_feature    <1. .>657
                     /gene="pol"

  Query Match             7.2%;  Score 540.1;  DB 1;  Length 657;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  565;  Conservative   60;  Mismatches   29;  Indels    7;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||| : |||||||| | |||| ||||||||||||: |||:|||||||
Db          1 TCCAAACCAAAGGCTCGGGTCTGCTCAGAACAGGCTAAATACTTAGGGGTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||||||||||:::|||||||| ||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
```

```
                 |||||||||||||||||||||:|::||| |||||:|| :|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                |||||||: |||:|||||||| |||:|||||:|||||:||||| ||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                |||||||||||||:|||||||||:|:|||||||:::||| |||||||||||||||  ||  |
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
                | |||  ||||||||:||||||||    ||||:|||||||||:|||||||||||| :||:|
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTA 357

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
                ||::|||||| |||||||||:||:|||||:|||||||||||||:|||:|||||: ||
Db        358 TATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAG 416

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
                | |||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:|||
Db        417 CCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCA 476

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
                |||||||: |||||||| ||||||||||||::||||||||||||||||||||| |||||||
Db        477 TTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACA 536

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
                || :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||
Db        537 GGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGG 596

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
                |||||||||||||||||||||:|||||||||||:|||||||||| | |||||:| ||
Db        597 AGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACC 656

Qy       3747 A 3747
                |
Db        657 A 657


RESULT 298
AB115229
LOCUS       AB115229                656 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-2.
ACCESSION   AB115229
VERSION     AB115229.1  GI:41016659
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP10-2"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
                     /note="HERV-W pol"
     repeat_region   1. .656
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 540;  DB 1;  Length 656;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  557;  Conservative   58;  Mismatches   33;  Indels    4;  Gaps    3;

Qy       3096 CSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGC 3155
                | |:|:| |||:|| |||| |||||||| ||||||||||:||||:|||||||||||||
Db          9 CAAAAGGTTCAGCTTTGCTTTCAGCAGGTTTAATACTTAGGGCTAAAATTATCCAAAGGC 68

Qy       3156 ACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACC 3215
                ||||:|||||||||||||||:::|||||||||||||||||||||||||||:||||||
Db         69 ACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCCAAAACT 128

Qy       3216 CTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAG 3275
                |||||||||||||||:|::||||||||:|||:|||:|||||  ||||:|||||||  |||||
Db         129 CTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT-CCCAG 187

Qy       3276 GTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATAC 3335
                ||: || :|||||| ||| :|||| :|||||:|||||||||||||  |||||| ||||||
Db         188 GTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGTCAATAC 247

Qy       3336 CCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCT 3395
                |||||||:||||||||:|:||||||:::||| |||||||||||||||  ||  | | |||
Db         248 CCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAGGCCCT 307

Qy       3396 TAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACA 3455
                ||||||||:|||||||||||||:|||| | :|||||||||||||:||:|||::||| |
Db         308 --AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATATGTCATA 365

Qy       3456 GAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACC 3515
                |||||| |||:||:|||||||:|||||||||||||:|||:|||||:||||||||||||
Db         366 -AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTTGCAACC 424

Qy       3516 YRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYG 3575
                ::||||::|||||| |||||||||:|||||||||||||||||||:|:||||||||||:|
Db         425 CGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATG 484

Qy       3576 GGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGA 3635
                |||||||| ||||||||||::|||||||||||||||||| ||||||||||||:|||||
Db         485 GGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGGAAGAGA 544

Qy       3636 TCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTG 3695
                | |||||||||||||||||||||:|||||::||||||| |||||||||||||||||||
Db         545 TTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGACTTGTG 604

Qy       3696 GCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                |||||||||||||:|||||||||||||:|||||||||||||||||||:|||||
Db         605 GCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 656


RESULT 299
AB115280
LOCUS       AB115280                655 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP8-3.
ACCESSION   AB115280
VERSION     AB115280.1  GI:41052508
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 655)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP8-3"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W pol"
     repeat_region   1. .655
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 539.9;  DB 1;  Length 655;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   30;  Indels    5;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:| |||:||||||| |||||||| |||||||||:||||:|||||||
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCA-RGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208
             |||||||||| :|||||||||||||||||:::|||||||||||||||||||||||||||:
Db         61 AAAGGCACCAGGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCC 120

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
             |||||| |||||||||||||:|::||||||||||:|||:|||:|||||:||||||||||
Db         121 CAAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGAT 180

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
             | |||||||: || :|||||| ||| :|||| :|||||:|||||||||||||  |||||
Db         181 T-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAG 239

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
              ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||
Db         240 TCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGA 299

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
             | | |||  ||||||||:|||||||||||||:|||| | :|||||||||||||:||:|||
Db         300 AGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATA 357

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
              ::||| | |||||| |||:||:||||||:||||||||||||||||:|||:|||||:|||||
Db         358 TGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCT 416

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::|||||| ||||||||:|||||||||||||||||||:|:||||||
Db         417 TGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATT 476

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||:|||||||||  |||||||||||::|||||||||||||||| ||||||||||||
Db         477 GTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGG 536

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |:|||||| |||||||||||||||||||||:|||||::|||||||  ||||||||||||
Db         537 GAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAG 596

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||:||||||||||||:||||||||||||||||||:|||||
Db         597 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 655


RESULT 300
AB051004
LOCUS       AB051004                 540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-7-1.
ACCESSION   AB051004
VERSION     AB051004.1  GI:14020829
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-7-1"
                     /db_xref="taxon:9606"
                     /chromosome="7"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47558.1"
                     /db_xref="GI:14020830"
                     /translation="NNFSTEINTTSVLVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQ
                     CIRWVTPPTQIVCLPSGIFFVCGTSAYRCLNGSSESMCFLSFLVPPMTIYTEQDLYSY
                     VISKPRNKRVPILPFVIGAGVLGALGTGIGGITTSTQFYYKLSQELNGDMERVADSLV
                     TLQDQLNSLAAVVLQNRRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 301
AB079501
LOCUS       AB079501                 540 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-1.
ACCESSION   AB079501
VERSION     AB079501.1  GI:22122075
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-1"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 302
AB114919
LOCUS       AB114919                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-5.
ACCESSION   AB114919
VERSION     AB114919.1  GI:41016427
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-5"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 303
AB114943
LOCUS       AB114943                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-8.
ACCESSION   AB114943
VERSION     AB114943.1  GI:41016451
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-8"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 304
AB114952
LOCUS       AB114952                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-4.
ACCESSION   AB114952
VERSION     AB114952.1  GI:41016460
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-4"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300
```

```
Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 305
AB114961
LOCUS AB114961 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE7-2.
ACCESSION AB114961
VERSION AB114961.1 GI:41016469
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE7-2"
/db_xref="taxon:9606"
/tissue_type="lung"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.2%; Score 538.4; DB 1; Length 540;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 536; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 306
AB114970
LOCUS AB114970 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE8-1.
ACCESSION AB114970
VERSION AB114970.1 GI:41016478
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE8-1"
/db_xref="taxon:9606"
/tissue_type="placenta"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.2%; Score 538.4; DB 1; Length 540;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 536; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 307
AB114973
LOCUS       AB114973                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-3.
ACCESSION   AB114973
VERSION     AB114973.1  GI:41016481
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-3"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 308
AB114989
LOCUS       AB114989                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-14.
ACCESSION   AB114989
VERSION     AB114989.1  GI:41016497
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-14"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 538.4;  DB 1;  Length 540;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 309
AB056490
LOCUS AB056490 654 bp DNA linear PRI 16-JAN-2002
DEFINITION Homo sapiens pol gene, HERV-W family, isolate:HWP6-2.
ACCESSION AB056490
VERSION AB056490.1 GI:15341218
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kim,H.S.
TITLE Sequence and phylogeny of HERV-W pol fragments
JOURNAL AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
PUBMED 11779355
REFERENCE 2 (bases 1 to 654)
AUTHORS Kim,H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .654
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="HWP6-2"
/db_xref="taxon:9606"
/chromosome="6"
/note="HERV-W pol
human endogenous retrovirus family :HERV-W"
gene 1. .654
/gene="pol"
misc_feature <1. .>654
/gene="pol"

Query Match 7.2%; Score 537.6; DB 1; Length 654;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 565; Conservative 57; Mismatches 31; Indels 6; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||| :|:|:|||||:||||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAATGAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||||||:|||||||||||||||||:::|||||||||| ||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTACACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db         121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCGGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:||| :|||||||||| :|||||:|||||:|||||||||||||||||||||||
Db         181 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||||||||||||||||  || |
Db         240 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:| ||||||||||:|||||| :|| ||||||||||:||:|||
Db         300 GGCCCT--AACCCAAGCCACAGTGTTAAGCTTGCCAAGGGGACAAGACTTTTCTTTATAA 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||||||||  ||:||:||||||:| |||||||| |||:|||:||||| ||||||
Db         358 GTCACAGAAAAAA--AGGAATAGCTCTAGTAGTCCTTACTCAGGTCCGAGGGAAGAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||||||:|:||| ||
Db         416 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCACTG 475

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| || |||||||||| ::|||||||||||||| ||||||||||||||||
Db         476 TTTACAGGTAGTGCTGACAGTAGCAGGCTTAGTATCTGAAGCCGTTAAAATAATACAGGG 535

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGT-GAAYRGCATACTCACTGCTAAAGGAG 3688
             :|||||||||||||||||||||||||||||:||  ||:  |||||||||||| ||||||||
Db         536 AAGAGATCTTACTGTGTGGACATCTCATGATGTAAAATCACATACTCACTGATAAAGGAG 595

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||: ||| |||||||:|||||||||||||||||:|||||
Db         596 ACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654
```

RESULT 310
AB115196
LOCUS AB115196 643 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-1.
ACCESSION AB115196
VERSION AB115196.1 GI:41016626
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 643)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .643
/organism="Homo sapiens"
/mol_type="other RNA"

```
                     /isolate="WP3-1"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .643
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 537.5;  DB 1;  Length 643;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  564;  Conservative   57;  Mismatches   22;  Indels   15;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             ||||||:|:|:| |||:|||||||| ||||||||| ||||||||||:||||:|||||||
Db             1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db            61 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db           121 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:||||:||||||||||:|||||:|||||:|||||||||||||||||||||||
Db           181 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||||:|:||||||:::|||  ||||||||| ||  || ||
Db           240 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGGCCTAAAGGA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:||||||||||||:||||||||: ||||||||||||:||:|||:
Db           300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGCGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||| ||:||:||||||:||||||||||||||:|||:| |||:||||||
Db           358 GTCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAAGGATGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||           :||||||:||||||||:||||| |||||||||||| ||:|:||||||
Db           417 GC----------CACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGCTGGCCTCATTG 465

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||||||||||||::|||||||||||||||||||||||||||||||||
Db           466 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 525

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|| ||| |||||||||||||| ||||||||:|||||::|| ||||||||||||||||||
Db           526 AAGGGATATTACTGTGTGGACGTCTCATGATGTGAATGGCGTACTCACTGCTAAAGGAGA 585

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||: ||||||||||:|||||||||||||| |||||:|||||
Db           586 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTATTTGAAGGGCCA 643


RESULT 311
AB115204
LOCUS       AB115204                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP5-3.
ACCESSION   AB115204
VERSION     AB115204.1  GI:41016634
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
    PUBMED  15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP5-3"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 537;  DB 1;  Length 650;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  563;  Conservative   58;  Mismatches   28;  Indels    8;  Gaps    4;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
             |||||| |:| |||||:||||||||||||||||||| |||||||:||||:||||||||
Db             2 CCAAACAAAATGCTCAGCTCTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATCCA 61

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
             ||||||| |:|||||||||||||||:::|||||||||||||||||||||||||||||:||
Db            62 AAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCA 121

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
             |||||||||||||||||:|::|||||||||:|||:| |:||||||||||:|||||||
Db           122 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGATT- 180

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
             |||||||:  ||:|||||| ||||:|||||:|||||:    ||||||||||||||||||
Db           181 CCCAGGTACAGCGAAATAGTCAGGCCATTATATACAC----TAAGGAAACTCAGAAAGCC 236

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
             |||||||||||:|||||||||:|:|||||||:::|||||||||| |||||||  ||  |
Db           237 AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAGAAG 296

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
             | |||  ||||||||:|||||||| ||||:||||||| :|||||||||||||:||:|||::
Db           297 GCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATATG 354

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
             || || ||| ||||||:||:|||||||:||||||||| ||||:|||:|||||: ||||||
Db           355 TCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTTG 413

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
             |||||::||||::|||||:|||||||:|||||||| ||||||||||||||:|:|||||||
Db           414 CAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTGT 473

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
             |||:|||||||||   ||||||||||::|||||||||||||||||||| ||||||||||:
Db           474 TTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGGA 533

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
             |||||||||||||||| ||||||||||:||||::|||||||||||||||||||||||||
Db           534 AGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGAC 593

Qy       3691 TTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||:|||||||||:|||||||||||||||||||| :|||||
Db           594 TTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 650


RESULT 312
AB056492
LOCUS       AB056492                657 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP7-2.
ACCESSION   AB056492
```

```
VERSION     AB056492.1  GI:15341220
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP7-2"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .657
                     /gene="pol"
     misc_feature    <1. .>657
                     /gene="pol"

  Query Match             7.2%;  Score 536.9;  DB 1;  Length 657;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  563;  Conservative   60;  Mismatches   31;  Indels    7;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCCAAACCAAAGCTCAACTTTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCCCCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTA 357

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
Db        358 TATGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAG 416

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
Db        417 CCTGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCA 476

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db        477 TTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACA 536

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db         537 GGAAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGG 596

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db         597 AGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACC 656

Qy       3747 A 3747
Db         657 A 657


RESULT 313
AB114922
LOCUS       AB114922                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-11.
ACCESSION   AB114922
VERSION     AB114922.1  GI:41016430
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-11"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```

```
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 314
AB114937
LOCUS       AB114937                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-13.
ACCESSION   AB114937
VERSION     AB114937.1  GI:41016445
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-13"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTAA 540


RESULT 315
AB114945
LOCUS       AB114945                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-14.
ACCESSION   AB114945
VERSION     AB114945.1  GI:41016453
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-14"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 316
AB114948
LOCUS AB114948 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE4-3.
ACCESSION AB114948
VERSION AB114948.1 GI:41016456
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE4-3"
/db_xref="taxon:9606"
/tissue_type="heart"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.2%; Score 536.8; DB 1; Length 540;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 535; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180
```

```
Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 317
AB114959
LOCUS AB114959 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE6-12.
ACCESSION AB114959
VERSION AB114959.1 GI:41016467
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE6-12"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.2%; Score 536.8; DB 1; Length 540;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 535; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
```

```
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 318
AB114966
LOCUS       AB114966                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-8.
ACCESSION   AB114966
VERSION     AB114966.1  GI:41016474
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-8"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 319
AB114968
LOCUS       AB114968                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-13.
ACCESSION   AB114968
VERSION     AB114968.1  GI:41016476
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-13"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60
```

```
Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||| ||||||||||||:|||:||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCGTTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 320
AB114978
LOCUS       AB114978                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-12.
ACCESSION   AB114978
VERSION     AB114978.1  GI:41016486
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-12"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 321
AB114979
LOCUS       AB114979                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-13.
ACCESSION   AB114979
VERSION     AB114979.1  GI:41016487
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-13"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.2%;  Score 536.8;  DB 1;  Length 540;
```

```
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAACCCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 322
AB056496
LOCUS       AB056496                656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP12-4.
ACCESSION   AB056496
VERSION     AB056496.1  GI:15341224
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP12-4"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.2%;  Score 536.5;  DB 1;  Length 656;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  562;  Conservative   60;  Mismatches   30;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
             | || : :|:|||||:|||||||||| | |||| ||||||||||||:||||:|||||||||
Db          2 CCAAACCAAAGCTCAACTCTGCTCAGAACAGGGTAAATACTTAGGGCTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
             ||||||||:|||||||||| |||||:::|||||||| ||||||||||||||||||:|||
Db         62 AGGCACCAGGGCCCTCAGTAAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
             ||||||||||||||||||:|::||| |||||:|| :|||:|||||||||||:||||||| |
Db        122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATAGCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
             ||||||: |||:|||||||| |||:|||||:|||||:|||||| ||||||||||||||||||
Db        181 CCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
             ||||||||||:|||||||||:|:||||||:::||| |||||||||||||   ||  | |
Db        241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
              |||  ||||||||:||||||||   ||||:||||||||:|||||||||||| :||:|||
Db        301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             ::|||||| |||||||||:||:||||||:|||||||||||||||:|||:||||||: |||
Db        359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::||||||:||||||||:|||||||||||||||||||:|:|||||
Db        418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||: |||||||| ||||||||||||::|||||||||||||||||| |||||||||
Db        478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
              :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:||||||||||:|||||||||| | ||||||:| |||
Db        598 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656


RESULT 323
AB056501
LOCUS       AB056501                656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP21-3.
ACCESSION   AB056501
VERSION     AB056501.1  GI:15341229
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP21-3"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.2%;  Score 536.5;  DB 1;  Length 656;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  562;  Conservative   60;  Mismatches   30;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
Db          2 CCAAACCAAAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
Db         62 AGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
Db        122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
Db        181 CCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
Db        241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
Db        301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
Db        359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
Db        418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
Db        478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
Db        538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        598 ACTTGTGGCTGTCAAACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656


RESULT 324
AB115203
LOCUS       AB115203                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP5-2.
ACCESSION   AB115203
VERSION     AB115203.1  GI:41016633
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP5-2"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 536.4;  DB 1;  Length 650;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  563;  Conservative   58;  Mismatches   29;  Indels    8;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACAAAATGCTCAGCTCTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACAGCGAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        236 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAGAA 295

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        296 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        354 GTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        413 GCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        473 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
```

```
Db        533 AAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:|||||||||||:||||||||||||||||| :|||||
Db        593 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 650


RESULT 325
AB115260
LOCUS       AB115260                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP3-4.
ACCESSION   AB115260
VERSION     AB115260.1  GI:41052488
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP3-4"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.2%;  Score 536.4;  DB 1;  Length 650;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  563;  Conservative   58;  Mismatches   29;  Indels    8;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              ||||||| |:| |||||:|||||||| |||||||||||| |||||||:||||:|||||||
Db          1 TCCAAACAAAATGCTCAGCTCTGCTCCCAGCAGGTTAAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||| |:||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||||||||||:|::|||||||||:|||:| |:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:  ||:|||||| ||||:|||||:|||||:    ||||||||||||||||||
Db        181 -CCCAGGTACAGCGAAATAGTCAGGCCATTATATATACAC----TAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::|||||||||| |||||||  ||  |
Db        236 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAGAA 295

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  |||||||:|||||||  ||||:||||||| :|||||||||||||:||:|||:
Db        296 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|| || ||| ||||||:||:||||||:|||||||||| ||||:|||:||||||: |||||

Db        354 GTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:||||||||:||||||| |||||||||||||:|:||||||
Db        413 GCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||   ||||||||||::|||||||||||||||||||| ||||||||||
Db        473 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||| ||||||||||:|||||::|||||||||||||||||||||||
Db        533 AAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:|||||||||||:||||||||||||||||| :|||||
Db        593 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 650


RESULT 326
AB096025
LOCUS       AB096025                544 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-5.
ACCESSION   AB096025
VERSION     AB096025.1  GI:30698501
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hong,K., Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of HERV-W env Fragments in Fetal Brain cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 544)
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-5"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .544
                     /gene="env"
     repeat_region   1. .544
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             7.2%;  Score 535.6;  DB 1;  Length 544;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  535;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGCTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||| |||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAAAGCTTTA 540

Qy       6739 GA 6740
              ||
Db        541 AA 542


RESULT 327
AB115000
LOCUS       AB115000                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-1.
ACCESSION   AB115000
VERSION     AB115000.1  GI:41016508
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-1"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.1%;  Score 535.2;  DB 1;  Length 540;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  534;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTCTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCCTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 328
AB056503
LOCUS       AB056503                 656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWPX-2.
ACCESSION   AB056503
VERSION     AB056503.1  GI:15341231
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWPX-2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.1%;  Score 534.9;  DB 1;  Length 656;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  561;  Conservative   60;  Mismatches   31;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
```

```
Db          2 CCAAACCAAAGCTCAACTCTGTTCAGAACAGGTTAAATACTTAGGGCTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211

Db         62 AGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271

Db        122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331

Db        181 CCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAAGTAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391

Db        241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448

Db        301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508

Db        359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568

Db        418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628

Db        478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688

Db        538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        598 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656
```

RESULT 329
AB115287
LOCUS AB115287 656 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP10-1.
ACCESSION AB115287
VERSION AB115287.1 GI:41052515
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 656)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .656
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWP10-1"
/db_xref="taxon:9606"
/cell_line="MCF7"
/note="HERV-W pol"
repeat_region 1. .656
/note="HERV-W pol"

Query Match 7.1%; Score 534.6; DB 1; Length 656;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 562; Conservative 59; Mismatches 33; Indels 6; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT--AARATTAT 3147

Db          1 TCCAAACCAAAAGTTCAGCTTTGCTCTCAGCAGGTTTAATACTTAGGGGCTAAAAATTAT 60

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207

Db         61 CCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATC 120

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267

Db        121 CCAAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGA 180

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327

Db        181 TT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAA 239

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387

Db        240 GTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAG 299

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447

Db        300 AAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTAT 357

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507

Db        358 ATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGC 416

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567

Db        417 TTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCAT 476

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627

Db        477 TGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAG 536

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687

Db        537 GGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGA 596

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        597 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 656
```

RESULT 330
AB096022
LOCUS AB096022 542 bp mRNA linear PRI 14-MAY-2003
DEFINITION Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-2.
ACCESSION AB096022
VERSION AB096022.1 GI:30698498
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hong,K., Shin,K., Yi,J. and Kim,H.
TITLE Isolation of HERV-W env Fragments in Fetal Brain cDNA library
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 542)

```
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-2"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .542
                     /gene="env"
     repeat_region   1. .542
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             7.1%;  Score 534;  DB 1;  Length 542;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  534;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||| |||| ||||||||||||||||||||| ||||||||||||||||||||||
Db        361 GCAGGAGTGCTGGGTGTACTAGGTACTGGCATTGGCGGCATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540

Qy       6739 GA 6740
              |
Db        541 AA 542


RESULT 331
AB056486
LOCUS       AB056486                654 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP4-1.
ACCESSION   AB056486
VERSION     AB056486.1  GI:15341214
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP4-1"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .654
                     /gene="pol"
     misc_feature    <1. .>654
                     /gene="pol"

  Query Match             7.1%;  Score 534;  DB 1;  Length 654;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  554;  Conservative   60;  Mismatches   40;  Indels    4;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||| ||:|:|:|||||:|||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAGACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             | |  |||||:||||||||||||||::|||||||||||||||||||||||| |||:|
Db         61 AGACACACCAGGGCCCTCAGTGAGGAATGCATCCAGCCTATACTGGCTTATCCTTATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||| |||||||| ||||:|::|||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCATAAAGCAATTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:  | :|||||||||||:|||||:| |||:|||||||||||||||||||||||
Db        181 -CCCAGGTACAGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:||||| |||:|:||||||:::|||||||||||||||||  ||  |
Db        240 GAATACCCATTTAGTAAGAAGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  || |||||:|| |||||||||:|||||||| :|||||||||||||:||:|| :
Db        300 GGCCCT--AATCCAAGCCCTAGTGTTAAGCTTGCCAAAGGGGCAAGACTTTTCTTTATGT 357

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :||||||| |||||||:||:||||||:||| ||||||||||:|||:|||||: || ||
Db        358 GTCACAGA-AAAAACAGGAATAGCTCTAGGATTCCTTACACAGGTCCGAGGGACCAGGTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||||
Db        417 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| |||  ||||||||||::|||||||||||||||||||||| ||||||||
Db        477 TTTATGGGTAATGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||  ||||||||:| | |::|||||||||||||||||||| |
Db        537 AAGAGATCTTACTGTGTGGATGTCTCATGATGGGGACGGCATACTCACTGCTAAAGGAAA 596
```

```
Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||  |||||||| |||||:|||| |||||||:|||||||||||||||||||:|||||
Db        597 CTTACGGCTGTCAAACAACCGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 332
AB114947
LOCUS       AB114947                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-17.
ACCESSION   AB114947
VERSION     AB114947.1  GI:41016455
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-17"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.1%;  Score 533.6;  DB 1;  Length 540;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  533;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCAGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| |||||:|||:|||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 333
A57089
LOCUS       A57089                  648 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 42 from Patent WO9628552.
ACCESSION   A57089
VERSION     A57089.1  GI:3713072
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 42 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              || ||||:||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
              ||||||||||||||||||:|||: ||||||||||||||||:|   ||::|||||||||||
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
              ||||   ||||||||| |||| :|||||:| |||||||  :|||| |||||||:|||||
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              |||||:|:|:|||||:|||||||||||||||||||||||| |||||:||||:||||||||
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||||:::|||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||::|| :|||||||||||:|||||:|||||:|||||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:| |||||:: ||   |||||||||| |  ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:||||||||||||:||||||||:|||||||||||||:||:|||::
```

```
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              |||||| |||||| ||:||:||||||:||||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||||::||||::| ||||:||||||:::||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 334
A57090
LOCUS       A57090                   648 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 43 from Patent WO9628552.
ACCESSION   A57090
VERSION     A57090.1  GI:3713073
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 43 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              || ||||:|||||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
              |||||||||||||||||:|||: |||||||||||||||:|   ||::||||||||||||
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
              ||||   |||||||| |||| :|||||:| |||||||  :|||| ||||||||:|||||
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              |||||:|:|:|||||:|||||||||||||||||||||||||| |||||:||||:||||||||
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||:::|||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||::|| :|||||||||||:|||||:|||||:|||||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:| ||||||:: ||    |||||||||||| |  ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482
```

```
Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:|||||||||||||:||||||||:|||||||||||||||:||:|||::
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              |||||| |||||| ||:||:||||||:||||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              ||||||::||||::| ||||:||||||:::||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 335
A80347
LOCUS       A80347                   648 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 42 from Patent EP0731168.
ACCESSION   A80347
VERSION     A80347.1  GI:6093074
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 42 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              || ||||:|||||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
              |||||||||||||||||:|||: |||||||||||||||:|   ||::||||||||||||
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
              ||||   |||||||| |||| :|||||:| |||||||  :|||| ||||||||:|||||
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              |||||:|:|:|||||:|||||||||||||||||||||||||| |||||:||||:||||||||
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||:::|||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||::|| :|||||||||||:|||||:|||||:|||||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:| ||||||:: ||    |||||||||||| |  ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
```

```
              | |||  ||||||||:||||||||||||:||||||||:|||||||||||||:||:|||::
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              |||||| |||||| ||:||:|||||:||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              |||||::||||::| ||||:||||||:::||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 336
A80348
LOCUS       A80348                   648 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 43 from Patent EP0731168.
ACCESSION   A80348
VERSION     A80348.1  GI:6093075
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 43 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.1%;  Score 533.4;  DB 1;  Length 648;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  548;  Conservative   63;  Mismatches   34;  Indels    4;  Gaps    3;

Qy       2911 GGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACC 2970
              || ||||:||||||||||||||||||||||||||||||||||||||:||:|::|||||||
Db          4 GGTATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCATACC 63

Qy       2971 TGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCT 3030
              ||||||||||||||||||:|||: |||||||||||||||:|   ||::||||||||||||
Db         64 TGGACACTCTTGTCCTTCAGTATATGGATGATTTACTTTTAGTGACCCATTCAGAAACCT 123

Qy       3031 TGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTT 3090
              ||||   |||||||| |||| :|||||:| |||||||  :|||| ||||||||:|||||
Db        124 TGTGATGTCAAGCCACACAAGTGCTCTTAACTTTCCTCTTTACCTCTGGCTACAAGGTTT 183

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              |||||:|:|:|||||:|||||||||||||||||||||||| |||||:||||:||||||||
Db        184 TCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATATTTAGGGCTAAAATTATCCA 243

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||:::|||||||||||||||||||||||| |||||:||
Db        244 AAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCTTCATCCCA 303

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              ||||||||||||||||||:|::|||||||||:|||:|||:|||||| |||:|||||||
Db        304 AAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCTGAATATGGATT- 362

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||::|| :|||||||||||:|||||:|||||:|||||||||||||||||||||||
Db        363 CCCAGGTTYGGTGAAATAGCCAGGCCATTAAATACACTAATTAAGGAAACTCAGAAAGCC 422

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:| |||||||:: ||   |||||||||| | ||  ||
Db        423 AATACCCATTTAGTAAGATGGACATCTGAAGCACAATCAGCTTTCCAGGCACTAAAGAAA 482

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:||||||||||||:||||||||:|||||||||||||:||:|||::
Db        483 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATATG 540
```

```
Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              |||||| |||||| ||:||:|||||:||||||||||||||:|| :|||||: ||||||
Db        541 TCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGACAAGCTTG 599

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
              |||||::||||::| ||||:||||||:::||||||||||||||||||
Db        600 CAACCCGTGGCATATCTGAGTAAGGARRCTGATGTAGTGGCAAAGGGTT 648


RESULT 337
AB115219
LOCUS       AB115219                 640 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP8-6.
ACCESSION   AB115219
VERSION     AB115219.1  GI:41016649
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 640)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP8-6"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
                     /note="HERV-W pol"
     repeat_region   1. .640
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 533.3;  DB 1;  Length 640;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   11;  Indels   21;  Gaps    5;

Qy       3093 AAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAA 3152
              || | |: :| |||:|| || || |||||||    |  |||||:||||:||||||||||
Db          5 AACCAAGGGGGTCAACTTTGGTCCCAGCAGGT----TTTTTAGGGCTAAAATTATCCAAA 60

Qy       3153 GGCACCA-RGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
              ||| ||| :|||||||||||||||:::|||||||||||||||||||||||||||||:|||
Db         61 GGCCCCAGGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCCAA 120

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
              |||||||||||||||||:|::|||||||||:|||:|||:|||||||||:||||||| |
Db        121 AACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT-C 179

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
              ||||||:||||:||||||||||:|||||:|||||:||||||||||||||||||||||||
Db        180 CCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGCCA 239

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
              ||||||||||:|||||||||:|:|||||||:::||||||||||||||||||
Db        240 ATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC--------- 290

Qy       3392 GCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRT 3451
              |||||||:||||||||||||:||||||||:|||||||||||||||:||:|||::|
Db        291 -----TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATATGT 345
```

```
Qy       3452 CACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGC 3511
Db        346 CACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTTGC 404

Qy       3512 AACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTT 3571
Db        405 AACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTGTT 464

Qy       3572 TAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRA 3631
Db        465 TACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAA 524

Qy       3632 GAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACT 3691
Db        525 GAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACT 584

Qy       3692 TGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        585 TGTGGCTGTCAGACAACTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCA 640
```

RESULT 338
AB115279
LOCUS AB115279 656 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP8-1.
ACCESSION AB115279
VERSION AB115279.1 GI:41052507
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 656)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .656
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWP8-1"
/db_xref="taxon:9606"
/cell_line="Hep G2"
/note="HERV-W pol"
repeat_region 1. .656
/note="HERV-W pol"

Query Match 7.1%; Score 532.9; DB 1; Length 656;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 561; Conservative 59; Mismatches 35; Indels 5; Gaps 4;

```
Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCA-GGTTAAATACTTAGGRCTAARATTAT 3147
Db          1 TCCAAACCCAAAAGGTCAGCTTTGCTCTCAGCAGGGTTTAATACTTAGGGCTAAAAATAT 60

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db         61 CCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATC 120

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db        121 CCAAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGA 180

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db        181 TT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAA 239

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db        240 GTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAG 299

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db        300 AAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTAT 357

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
Db        358 ATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGC 416

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
Db        417 TTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCAT 476

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
Db        477 TGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAG 536

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
Db        537 GGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGA 596

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 GACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 656
```

RESULT 339
AB115212
LOCUS AB115212 649 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-5.
ACCESSION AB115212
VERSION AB115212.1 GI:41016642
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 649)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .649
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WP7-5"
/db_xref="taxon:9606"
/tissue_type="lung"
/note="HERV-W pol"
repeat_region 1. .649
/note="HERV-W pol"

Query Match 7.1%; Score 532.1; DB 1; Length 649;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 561; Conservative 57; Mismatches 31; Indels 9; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        472 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        532 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        592 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 649
```

RESULT 340
AB114930
LOCUS AB114930 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE1-19.
ACCESSION AB114930
VERSION AB114930.1 GI:41016438
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE1-19"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.1%; Score 532; DB 1; Length 540;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 532; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 341
AB115300
LOCUS AB115300 653 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP15-6.
ACCESSION AB115300
VERSION AB115300.1 GI:41052528
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 653)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,

```
           Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP15-6"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W pol"
     repeat_region   1. .653
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 531.7;  DB 1;  Length 653;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  566;  Conservative   59;  Mismatches   27;  Indels    7;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTA-GGRCTAARATTATC 3148
             |||||||:|:|:|||||:|||||||||||||||||||||||||||| ||:||||:||| ||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGGCTAAAATTGTC 60

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208
             ||||||||||:|||||||||||| ||:::|||||||||||||||||||||||| |||:
Db         61 CAAAGGCACCAGGGCCCTCAGTGAGAAATGTATCCAGCCTATACTGGCTTATCCTTATCC 120

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
             ||||||  |||||||||||||:| :||||||||:|||:|||:||||||||||:||||||
Db         121 CAAAACACTAAAGCAACTAAGAGCGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGAT 180

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
             | |||||||:  | :|||||||||||:|||||:|||||:|||||| | ||||||||||||
Db         181 T-CCCAGGTACAGTGAAATAGCCAGGTCATTATATACACTAATTAGGAAAACTCAGAAAG 239

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
             ||||| ||||||||:|||||||||:|:||||||:::|||||||||||||| ||||||||
Db         240 CCAATGCCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGG-CCCTAAAG 298

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
             |||||| |||||| |  :||    ||||||:||||||| :||||||||||||:||:|||
Db         299 AAGGCCCTAAACCAAGCCCC---AGTTAAGTTTGCCAATGGGGCAAGACTTTTCTTTATA 355

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             ::|||||| |||||||||:||:||||| :|||| |||||||| |:|| :|||||:|||||
Db         356 TGTCACAG-AAAAAACAGGAATAGCTCCAGGAGGCCTTACACCGGTCTGAGGGATGAGCT 414

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::||||||:||||||||:|||||||||||| ||||||||:|:|||||
Db         415 TGCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTAGTGGCCAAGGGTTGGCCTCATT 474

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||:|||||||||  |||||||||||::|||||||||||||||||| ||||||||||
Db         475 GTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGG 534

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |:|||||| |||||||||||||||||||||:||||::||||||| ||||||||||||
Db         535 GAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAG 594

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||:|||||||||||:|||||||||||||||||||:|||||
Db         595 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 653


RESULT 342
AB115232
LOCUS       AB115232                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-8.
ACCESSION   AB115232
VERSION     AB115232.1  GI:41016662
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP10-8"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 531.6;  DB 1;  Length 650;
  Best Local Similarity   85.1%;  Pred. No. 0;
  Matches  560;  Conservative   58;  Mismatches   32;  Indels    8;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             ||||| | |:| |||||:| |||||||||||||||||| |||||||:||||:|||||||
Db          1 TCCAATCAAAATGCTCAGCACTGCTCACAGCAGGTTAAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||||| |:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||:|::|||||||||:|||:| |:||||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||:  ||:|||||| |||||||||:| |||:    ||||||||||||||||||
Db         181 -CCCAGGTACAGCGAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::|||||||||| |||||||  ||  |
Db         236 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAGAA 295

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||||||:|||||| ||||:|||||||  :|||||||||||||:||:|||:
Db         296 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|| || ||| |||||||:||:|||||:|||||||||| ||||:|||:|||||: |||||
Db         354 GTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             |||||::||||::||||||:||||||||:||||||| |||||||||||||:|:|||||||
Db         413 GCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  |||||||||::|||||||||||||||||||| ||||||||||
Db         473 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||| ||||||||||:|||||::|||||||||||||||||||||||
Db         533 AAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||:|||||||||||:|||||||||||||||||||| :|||||
Db         593 TTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 650
```

```
RESULT 343
AB126119
LOCUS       AB126119                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-7.
ACCESSION   AB126119
VERSION     AB126119.1  GI:47169719
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-7"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 531.6;  DB 1;  Length 650;
  Best Local Similarity   86.2%;  Pred. No. 0;
  Matches  567;  Conservative   57;  Mismatches   26;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||| |||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||| |||||||| |:: ||||||| :|||:|||:||||||| ||:|||||||
Db        121 AAAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:||||| ||||||| |||||||||||||
Db        180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::||| ||||||||||||||  ||  |
Db        239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||||:||||||||:|||||||||||     :|||:
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||| | |||||||||:||:||||||:||||||||||||||||:|||:||||||: |||||
Db        354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATAAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:|||| |||:|||||||||||||||||:|:||||||
Db        413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||| ||||||||||||||||

Db        473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :||| |||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        533 AAGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||:||||||||||:|||||||  ||||||||||:|||||
Db        593 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 344
AB126113
LOCUS       AB126113                642 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-9.
ACCESSION   AB126113
VERSION     AB126113.1  GI:47169713
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 642)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-9"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .642
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 531.2;  DB 1;  Length 642;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  553;  Conservative   58;  Mismatches   31;  Indels   16;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:|::|| ||||||:|||:|||:|||||||||| |||||||
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTTCTTGGCATAACAGGTTTCTGCCGAACATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                | |||: |||:|||||||||| :|||||:|||||:|||||||||||||||||| |||
Db        181 -GACGGGTACGGCGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAGAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||| || ||||:|||||||||:|:|||||:::||| |||||||||||||
Db        240 CAATATCCTTTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCC------- 292

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                     ||||||||:||| |||||||||:|| |||||:||||||||||||:||:|||:
Db        293 -------TAACCCAAGCCCCCGTGTTAAGCTTCCCAACGGGGCAAGACTTTTCTTTATAT 345
```

```
Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||| | :|||||||:|||||||||||||||:|||:|||||: ||| |
Db        346 GTCACAG-AAAAAACAGCACTAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACCAGCCT 404

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||:: |||||:||||||||:|||||||||||||||||||  :|:||||||
Db        405 GCAACCCGTGGCATCCCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTCAGCCTCATTG 464

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  ||| |||| ||::|||||||  ||| ||||||||||||||||||
Db        465 TTTACGGGTAGTGGCAGCAATAGCCGTCTTAGTATCTATAGCGGTTAAAATAATACAGGG 524

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||| |||:|||||::|||||||||||||||||||||
Db        525 AAGAGATCTTACTGTGTGGACATCTCGTGATGTGAATGGCATACTCACTGCTAAAGGAGA 584

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||| |||||||:||||||||||||:| ||||||||||||||||||:|||||
Db        585 CTTGTGGCTGTTAGACAACTGTTTACTTAAATATTAGGCTCTATTACTTGAAGGGCCA 642


RESULT 345
AB056484
LOCUS       AB056484                649 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP3-2.
ACCESSION   AB056484
VERSION     AB056484.1  GI:15341212
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP3-2"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .649
                     /gene="pol"
     misc_feature    <1. .>649
                     /gene="pol"

  Query Match             7.1%;  Score 530.9;  DB 1;  Length 649;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  561;  Conservative   56;  Mismatches   32;  Indels   9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||| |||||||||||||||| |:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCCCAGCAGGTTAAATACTTACGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||:|::|||||     |||: ||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAATAGGGTTCCT-----TAACGGGTTTCTGCCGAATATGGATT 175

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: || :||||| |||| |||| :|||||:|||||||||||||||||||||||
Db        176 -CCCAGGTACGGTGAAATACCCAGAACATTCTATACACTAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||: :|||| |||||| ||| ||  |   |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAAGCGGAAGTAGCTTTCTAGGACCTAACGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||||||:|||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:|||||||||||||||||:|||:|||||:|||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | |||||||||::|||||||||||  |||||||| ||||||||
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:|||||||||||:|||||||||||||||||||:| |||
Db        592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 346
AB126115
LOCUS       AB126115                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-3.
ACCESSION   AB126115
VERSION     AB126115.1  GI:47169715
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-3"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 530.6;  DB 1;  Length 650;
  Best Local Similarity   86.1%;  Pred. No. 0;
```

```
  Matches  566;  Conservative   57;  Mismatches   26;  Indels    8;  Gaps    5;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
Db          2 CCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCTTAAATACTTAGGGCTAAAATTATCCA 61

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
Db            62 AAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCA 121

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
Db           122 AAACCTTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT- 179

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
Db           180 CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGCC 239

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
Db           240 AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAG 299

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
Db           300 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATATG 354

Qy       3451 TCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
Db           355 TCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTTG 413

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
Db           414 CAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 473

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
Db           474 TTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGGA 533

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
Db           534 AGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGAC 593

Qy       3691 TTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db           594 TTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 347
AB056497
LOCUS       AB056497                656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP13-1.
ACCESSION   AB056497
VERSION     AB056497.1  GI:15341225
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP13-1"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.1%;  Score 530.5;  DB 1;  Length 656;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  559;  Conservative   59;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
Db             2 CCAAACCAAAGCTCAACTCTGCTCAGAACAGGCTAAATATTTAGGGCTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
Db            62 AGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
Db           122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
Db           181 CCATGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
Db           241 ATACCCATTTAGTAAGATGGGCACCTGAAGCTGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
Db           301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGAGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
Db           359 TGTCACAGA-AAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
Db           418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
Db           478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
Db           538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db           598 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656


RESULT 348
AB114991
LOCUS       AB114991                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-2.
ACCESSION   AB114991
VERSION     AB114991.1  GI:41016499
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-2"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.1%;  Score 530.4;  DB 1;  Length 540;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 349
AB056494
LOCUS       AB056494                656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP10-3.
ACCESSION   AB056494
VERSION     AB056494.1  GI:15341222
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP10-3"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.1%;  Score 530.1;  DB 1;  Length 656;
  Best Local Similarity   84.7%;  Pred. No. 0;
  Matches  558;  Conservative   60;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
             | || : :|:|||||:||||| ||| | |||||||||||||||||: |||:||||||||| 
Db          2 CCAAACCAAAGCTCAACTCTGTTCAGAACAGGTTAAATACTTAGGGTTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
             ||||||||:|| |||||||||||||:::|||||||| |||||||||||| |||||:|||
Db         62 AGGCACCAGGGGCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCTTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
             ||||||||||||||||||:|::||| ||||||:|| :|||:|||||| |||:||||||| |
Db        122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCGGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
             ||||||: |||:|||||||| |||:|||||:|||||:||||| ||||||||||||||||||
Db        181 CCAGGTACGGCGAAATAGCTAGGCCATTATATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
             ||||||||||:||||||||||:|:||||||:::||| |||||||||||||  || | |
Db        241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
             |||  |||||||:|||||||    ||||:||||||||:|||||||||||| :||:|||
Db        301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             ::|||||| |||||||||:||:||||||:|||||||||||| |||:|||:|||||: |||
Db        359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACGCAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::|||||||:|||||||:|||||||||||||||||||||:|:|||||
Db        418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||: |||||||| |||||||||||::|||||||||||||||||||| |||||||||
Db        478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597
```

```
Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||||:||||||||||:|||||||||| | ||||||:| |||
Db        598 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656


RESULT 350
AB126112
LOCUS       AB126112                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-7.
ACCESSION   AB126112
VERSION     AB126112.1  GI:47169712
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-7"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 530;  DB 1;  Length 650;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  566;  Conservative   57;  Mismatches   27;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||| ||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||| |||||||| |:: ||||||| :|||:|||:|||||||| ||:|||||||
Db         121 AAAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:|||| ||||||| ||||||||||||||
Db         180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db         239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||:||||||||:|||||||||     :|||:
Db         299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||| | |||||||||:||:||||||:||||||||||||||:|||:|||||: |||||
Db         354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATAAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:|||| |||:|||||||||||||||||||:|:||||||
Db         413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||| ||||||||||||||||
Db         473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :||| ||||||||||||||||||||||:||||||::|||||| ||||||||||||||||
Db         533 AAGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACCCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:||||||||||:|||||||  ||||||||||:|||||
Db         593 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 351
AB126114
LOCUS       AB126114                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-2.
ACCESSION   AB126114
VERSION     AB126114.1  GI:47169714
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-2"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 530;  DB 1;  Length 650;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  566;  Conservative   57;  Mismatches   27;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||| ||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||||  :||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCGGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||| |||||||| |:: ||||||| :|||:|||:|||||||| ||:|||||||
Db         121 AAAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:|||| ||||||| ||||||||||||||
Db         180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
```

```
Db        239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:||||||||:||||||||||     :|||:
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||| | |||||||||:||:|||||||:|||||||||||||||:|||:|||||:||||||
Db        354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::|||||||:|||| |||:|||||||||||||||||||:|:||||||
Db        413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||| ||||||||||||||||
Db        473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :||| ||||||||||||||||||||||:||||::|||||||||||||||||||||||||
Db        533 AAGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||:||||||||||||:||||||| ||||||||||:|||||
Db        593 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 352
AB126116
LOCUS       AB126116                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-4.
ACCESSION   AB126116
VERSION     AB126116.1  GI:47169716
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-4"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 530;  DB 1;  Length 650;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  566;  Conservative   57;  Mismatches   27;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||  ||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||| |||||||| |:: ||||||| :|||:|||:|||||||  ||:|||||||
Db         121 AAAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:||||| ||||||| ||||||||||||||
Db         180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||||  ||  |
Db         239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:||||||||:||||||||||     :|||:
Db         299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||| | |||||||||:||:|||||||:|||||||||||||||:|||:|||||:||||||
Db         354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::|||||||:|||| |||:|||||||||||||||||||:|:||||||
Db         413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||:|||||||||  ||||||||||::|||||||||||||| ||||||||||||||||
Db         473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :||| ||||||||||||||||||||||:||||::|||||||||||||||||||||||||
Db         533 AAGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||||:||||||||||||:||||||| ||||||||||:|||||
Db         593 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 353
BV491951/c
LOCUS       BV491951                667 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S221P61014FG7.T0 Yvonne Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV491951
VERSION     BV491951.1  GI:62335621
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 667
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
```

(Donald,Karlien,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

```
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Yvonne"
     STS             <1. .>667
```

Query Match 7.1%; Score 529.9; DB 1; Length 667;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 575; Conservative 33; Mismatches 56; Indels 5; Gaps 3;

```
Qy       3872 TAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTA 3931
Db        667 TAGAGGTTCCCTTGACTGATTCCGA-CTTCAACTTGTACACTGATGGAAGTTTATTTTTA 609

Qy       3932 GAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGT 3991
Db        608 GAAAAAGGACTTTAAAAAGCGTATTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGT 549

Qy       3992 AATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCA 4051
Db        548 AATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAATAATAGCCCTCACTCAGGTG 489

Qy       4052 CTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTA 4111
Db        488 CTAGAATTAGGAGAAAGGAAAAGGGTAAATATATATACAGACTCTAAATATGCTTACCTA 429

Qy       4112 GTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACA 4171
Db        428 GTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAAGGACAA 369

Qy       4172 CCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAG 4231
Db        368 CCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGCACAGAAATCTAAAGAG 309

Qy       4232 GTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAY 4291
Db        308 GTGGCAATCTTACACTACTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAAC 249

Qy       4292 TGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
Db        248 CGCCAAGTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACCCTCCATAAGAAATGCTT 189
```

```
Qy       4352 ATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCA 4411
Db        188 ATAGAAGGACCCCTAGTATGGGGTAATCCCCTCCAGGAAACCAAGCCCCAGTACTCAGCA 129

Qy       4412 GGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCC 4468
Db        128 AGAGAGAGAGAATGGGGAACCTCACAAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCC 69

Qy       4469 ACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAACCCTT 4528
Db          68 ACTGAAGAAGGGAAAATACTTTTGTCTGCAGCTAACCCATGGAAATTA-TTTAAACCCTT 10

Qy       4529 CATCAAACC 4537
Db           9 CACCAAACC 1
```

```
RESULT 354
AB056487
LOCUS       AB056487                654 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP4-2.
ACCESSION   AB056487
VERSION     AB056487.1  GI:15341215
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP4-2"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .654
                     /gene="pol"
     misc_feature    <1. .>654
                     /gene="pol"
```

Query Match 7.1%; Score 529.2; DB 1; Length 654;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 551; Conservative 60; Mismatches 43; Indels 4; Gaps 3;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAGACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db          61 AGACACACCAGGGCCCTCAGTGAGGAATGCATCCAGCCTATACTGGCTTATCCTTATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db         121 AAAACCATAAAGCAATTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
```

```
Db        181 -CCCAGGTACAGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db        240 GAATACCCATCTAGTAAGAAGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db        300 GGCCCT--AATCCAAGCCCTAGTGTTAAGCTTGCCAAAGGGGCAAGACTTTTCTTTACGT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db        358 GTCACAG-AAAAAACAGGAATAGCTCTAGGATTCCTTACACAGGTCCGAGGGACCAGGTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db        417 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db        477 TTTATGGGTAATGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689

Db        537 AAGAGATCTTACTGTGTGGATGTCTCATGATGTGGACGGCATACTCACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        597 CTTACCGCTGTCAGACGACCCGTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 355
AB056485
LOCUS       AB056485                649 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP3-3.
ACCESSION   AB056485
VERSION     AB056485.1  GI:15341213
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP3-3"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .649
                     /gene="pol"
     misc_feature    <1. .>649
                     /gene="pol"

  Query Match             7.1%;  Score 528.9;  DB 1;  Length 649;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches  559;  Conservative   57;  Mismatches   33;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149

Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209

Db          61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269

Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329

Db         181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db         295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db         353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db         412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db         472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689

Db         532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db         592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 356
AB114981
LOCUS       AB114981                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-3.
ACCESSION   AB114981
VERSION     AB114981.1  GI:41016489
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
```

```
                     /mol_type="other RNA"
                     /isolate="WE10-3"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.1%;  Score 528.8;  DB 1;  Length 540;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  530;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||:|||:|||||||||||||||||| ||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||||||||
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 357
AB115286
LOCUS       AB115286                649 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP9-4.
ACCESSION   AB115286
VERSION     AB115286.1  GI:41052514
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)

FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP9-4"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W pol"
     repeat_region   1. .649
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 528.5;  DB 1;  Length 649;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  558;  Conservative   58;  Mismatches   33;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:||||:|||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| ||||||||||:::|||||||||||||||||| |||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: |||||||:|||:|||:||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||:|||||||||  |||||:|     |||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::| |  ||||||||||||   || |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAGGCAGCTTTCCAGGCCGTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                 |||  ||||||||:||||||||||||:|||||| |:|||||||||||||:||: ||:
Db        295 CACCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAGCGGGGCAAGACTTTTCTTTCTAT 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||||||||  ||:||:||||||:||||||||||||||:|||:|||||: |||||
Db        353 GTCACAGAAAAAAT-AGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACCAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||::||||||:||||||||:|||||||||||||||||||:|:||||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||| |||||||||||::|||||||||||||||||||| ||||||||
Db        472 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||| ||||||||:|||||::|||||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACGTCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||| |||: |||||||||||:|||||  |||||||||||||:|||||
Db        592 CTTGTGGCTGTCAGATAACCATTTACTTAAATATCAGGTTCTATTACTTGAAGGGCCA 649


RESULT 358
AB126124
LOCUS       AB126124                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP7-5.
ACCESSION   AB126124
VERSION     AB126124.1  GI:47169724
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
```

```
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP7-5"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.1%;  Score 528.4;  DB 1;  Length 650;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  565;  Conservative   57;  Mismatches   28;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:||||||:|||||||||||||||  ||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCCTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||| ||||||||  |:: |||||||  :|||:|||:|||||||  ||:|||||||
Db         121 AAAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:|||||  |||||||  ||||||||||||||
Db         180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||:|:||||||:::||| |||||||||||||  ||  |
Db         239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  |||||||:||||||||||:||||||||:|||||||||||     :|||:
Db         299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||| | ||||||||||:||:|||||:|||||||||||||||:|||:|||||: |||||
Db         354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATAAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:|||| |||:|||||||||||||||||||:|:||||||
Db         413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  |||||||||||::|||||||||||||| ||||||||||||||||
Db         473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :||| |||||||||||||||||||||||||:||||||::|||||||||||||||||||||||
Db         533 AAGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:||||||||||||:|||||||  ||||||||||:|||||
Db         593 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTGAATTACTTGAAGGGCCA 650


RESULT 359
AB126107
```

```
LOCUS       AB126107                642 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-1.
ACCESSION   AB126107
VERSION     AB126107.1  GI:47169707
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 642)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-1"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .642
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 528;  DB 1;  Length 642;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches  551;  Conservative   58;  Mismatches   33;  Indels   16;  Gaps    3;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|: :|||||:||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAACGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||:|::|| ||||||:|||:|||:|||||||||| |||||||
Db         121 AAAACCCTAAAGCAACTAAGAGGGTTTCTTGGCATAACAGGTTTCTGCCGAACATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              | |||: |||:|||||||||| :|||||:|||||:|||||||||||||||||| |||
Db         181 -GACGGGTACGGCGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAGAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||| || ||||:|||||||||:|:|||| |:::||| ||||||||||||||
Db         240 CAATATCCTTTTAGTAAGATGGACACCTGAGGCAGAAGCGGCTTTCCAGGCCC------- 292

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                      ||||||||:||| |||||||:|| |||||:||||||||||||:||:|||:
Db         293 -------TAACCCAAGCCCCCGTGTTAAGCTTCCCAACGGGGCAAGACTTTTCTTTATAT 345

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||| | :|||||:|||||||||||||||:|||:|||||: ||| |
Db         346 GTCACAG-AAAAAACAGCACTAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACCAGCCT 404

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||:: |||||:||||||||:|||||||||||||||||||  :|:||||||
Db         405 GCAACCCGTGGCATCCCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTCAGCCTCATTG 464

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  ||| |||| ||::|||||||  ||| ||||||||||||||||||
Db         465 TTTACGGGTAGTGGCAGCAATAGCCGTCTTAGTATCTATAGCGGTTAAAATAATACAGGG 524
```

```
Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        525 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 584

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||| |||||||:| ||||||||||:| ||||||||||||||||||:|||||
Db        585 CTTGTGGCTGTTAGACAACTGGTTACTTAAATATTAGGCTCTATTACTTGAAGGGCCA 642


RESULT 360
AB079509
LOCUS       AB079509                 541 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-9.
ACCESSION   AB079509
VERSION     AB079509.1  GI:22122083
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-9"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .541
                     /rpt_family="HERV-W"

  Query Match             7.0%;  Score 527.9;  DB 1;  Length 541;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGC-TTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
              ||||||:|||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATA 300

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 360

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 420

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        481 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

Qy       6738 A 6738
              |
Db        541 A 541


RESULT 361
AB126120
LOCUS       AB126120                 650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-8.
ACCESSION   AB126120
VERSION     AB126120.1  GI:47169720
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-8"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 527.4;  DB 1;  Length 650;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  564;  Conservative   57;  Mismatches   28;  Indels    8;  Gaps    5;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              ||||||:|:|:|||||:|||||||||||||  | ||||||||||:||||:||||||||
Db          2 CCAAACCAAAGGCTCAGCTCTGCTCACAGCAGCCTGAATACTTAGGGCTAAAATTATCCA 61

Qy       3151 AAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCA 3210
              |||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:||
Db         62 AAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCA 121

Qy       3211 AAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTC 3270
              |||||||||| |||||||| |:: ||||||| :|||:|||:||||||| ||:|||||||
Db        122 AAACCCTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT- 179

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||:||| :|||||||||| :|||||:|||| ||||||| ||||||||||||||||
Db        180 CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGCC 239

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        240 AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAG 299
```

```
Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
             | |||  |||||||:|||||||||||||:||||||||:||||||||||     :|||::
Db        300 GCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATATG 354

Qy       3451 TCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
             |||| | |||||||||:||:||||||:|||||||||||||||||:|||:|||||: ||||||
Db        355 TCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATAAGCTTG 413

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
             |||||::||||::||||||:|||| |||:|||||||||||||||||||:|:|||||||
Db        414 CAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 473

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
             |||:|||||||||  ||||||||||::|||||||||||||| ||||||||||||||||:
Db        474 TTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGGA 533

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
             ||| |||||||||||||||||||||:|||||::|||||||||||||||||||||||||||
Db        534 AGATATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGAC 593

Qy       3691 TTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||:||||||||||:||||||| ||||||||||:|||||
Db        594 TTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 362
AB056483
LOCUS       AB056483               649 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP3-1.
ACCESSION   AB056483
VERSION     AB056483.1  GI:15341211
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP3-1"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .649
                     /gene="pol"
     misc_feature    <1. .>649
                     /gene="pol"

  Query Match             7.0%;  Score 527.3;  DB 1;  Length 649;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  558;  Conservative   57;  Mismatches   34;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||| |||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db         61 AAAGGCGCCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: |||||||:|||:|||:|||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||
Db         181 -CCCAGGTACGGCGAAATAGCCAGAGCATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||||:|:|||| |:::|||  |||||||||| |  || |
Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||| ||||||:|||||||:|| ||||||||||:||:|||
Db         295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:|||||||||||||||:|||:|||||:||||||
Db         353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db         412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||| ||| | |||||||||::|||||||||| |||||||| ||||||||
Db         472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db         532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:||||||||||||:|||||||||||||||||||||:| |||
Db         592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 363
AB115268
LOCUS       AB115268               649 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP5-9.
ACCESSION   AB115268
VERSION     AB115268.1  GI:41052496
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP5-9"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W pol"
```

```
     repeat_region   1. .649
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 527.3;  DB 1;  Length 649;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  558;  Conservative   57;  Mismatches   34;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: |||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||||  |||||:|     |||||||||||||||||||||||
Db         181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||||  |  ||  |
Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||  |||||:|||||||||:||  ||||||||||:||:|||
Db         295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:||||||||||||||||:|||:||||||:||||||
Db         353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::|||||||:|||||||:|||||||||||||||| ||||:|: |||||
Db         412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | ||||||||||::|||||||||  |||||||| ||||||||
Db         472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db         532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGG 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||:||||||||||||:||||||||||||||||||:| |||
Db         592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 364
AB115289
LOCUS       AB115289                649 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP10-4.
ACCESSION   AB115289
VERSION     AB115289.1  GI:41052517
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP10-4"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W pol"
     repeat_region   1. .649
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 527.3;  DB 1;  Length 649;
  Best Local Similarity   84.8%;  Pred. No. 0;
  Matches  558;  Conservative   57;  Mismatches   34;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: |||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||||  |||||:|     |||||||||||||||||||||||
Db         181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||||  |  ||  |
Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||  |||||:|||||||||:||  ||||||||||:||:|||
Db         295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||| | ||||||||:||:||||||:||||||||||||||||:|||:||||||:||||||
Db         353 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::|||||||:|||||||:|||||||||||||||| ||||:|: |||||
Db         412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | ||||||||||::|||||||||  |||||||| ||||||||
Db         472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db         532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||:||||||||||||:||||||||||||||||||:| |||
Db         592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 365
AB056504
LOCUS       AB056504                656 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWPY-3.
ACCESSION   AB056504
VERSION     AB056504.1  GI:15341232
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWPY-3"
                     /db_xref="taxon:9606"
                     /chromosome="Y"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .656
                     /gene="pol"
     misc_feature    <1. .>656
                     /gene="pol"

  Query Match             7.0%;  Score 527.3;  DB 1;  Length 656;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  557;  Conservative   59;  Mismatches   36;  Indels    7;  Gaps    4;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
              | || : :|:| |||:|||||||| | ||||  ||| |||||||:|| |:|||||||||
Db          2 CCAAACCAAAGGTCAACTCTGCTCAGAACAGGGCAAAAACTTAGGGCTGAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
              ||||||||:|||||||||||||||||| ::|||||||| ||||||| ||||||||||:|||
Db         62 AGGCACCAGGGCCCTCAGTGAGGAAGGTATCCAGCCCATACTGGGTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
              |||| |||||||||||||:|::||| |||||:|| :|||:||||||||||:||||||| |
Db        122 AACCGTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
              ||||||: |||:|||||||| |||:|||||:|||||:||||| |||||||||||||||||
Db        181 CCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
              |||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  | |
Db        241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
              |||  ||||||||:||||||||   ||||:|||||||:|||||||||||| :||:|||
Db        301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
              ::|||||| |||||||||:||:|||||:|||||||||||||||:|||:|||||: |||
Db        359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
              |||||||::||||::||||||:||||||||:|||||||||||||||||||:|:|||||
Db        418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
              |||||: |||||||| ||||||||||||::|||||||||||||||||||| |||||||||
Db        478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
              :|||||||||||||||||||||||||||:||||||::|||||||||||||||||||||||
Db        538 AAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 597

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||:||||||||||||:|||||||||| | |||||||:| |||
Db        598 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 656


RESULT 366
AB115200
LOCUS       AB115200                636 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-8.
ACCESSION   AB115200
VERSION     AB115200.1  GI:41016630
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 636)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP3-8"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .636
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 527.2;  DB 1;  Length 636;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  567;  Conservative   60;  Mismatches    8;  Indels   22;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||     |||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGT----TACTTAGGGCTAAAATTATCC 56

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:|
Db         57 AAAGGCACCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCC 116

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||:|::||||||||:|||:|||:||||||||||:|||||||
Db         117 AAAACCCTAAAGCAACTAAGGGGATTCCTTGGCGTAATAGGTTTCTGCCGAAAATGGATT 176

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||||:||||||||||:|||||:|||||:||||||||||||||||||||||
Db         177 -CCCAGGTATGGCGAAATAGCCAGGTCATTAAATACACTAATTAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||||:||||||||:::|||||||||||||||||||||||||||||||
Db         236 CAATACCCATTTAGTAAGATGGACAACTGAAGTAGAAGTGGCTTTCCAGGCCC------- 288

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                      ||||||||:||||||||||||:||||||||:|||||||||||||||:||:|||:
Db         289 -------TAACCCAAGCCCCAGTGTTAAGTTTGCCAACAGGGCAAGACTTTTCTTCATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:|||||:|||||||||||||||:|||:|||||:||||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCGAGGGATGAGCTT 400
```

```
Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||||
Db        401 GCAACCTGTGGCATACCTGACTAAGGAAATTGATGTAGTGGCAAAGGGTTGACCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||||||||||||||||||::|||||||||||||||||||||| ||||||
Db        461 TTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAA-ACAGGG 519

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :||||||||||||| ||||||||||||||:|||||::|| |||||||||||||||| |
Db        520 AAGAGATCTTACTGGGTGGACATCTCATGATGTGAATGGCCTACTCACTGCTAAAGGAAA 579

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||| ||||| || |||||:| ||||||||||:||||||||| |||||||||:||||
Db        580 CTTGGGGCTGGCAAACAACTGGTTACTTAAATGTCAGGCTCT-TTACTTGAAGGGCC 635


RESULT 367
AB126117
LOCUS       AB126117                650 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-5.
ACCESSION   AB126117
VERSION     AB126117.1  GI:47169717
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-5"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 526.8;  DB 1;  Length 650;
  Best Local Similarity   85.7%;  Pred. No. 0;
  Matches  564;  Conservative   57;  Mismatches   29;  Indels    8;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|| |||||||||||| |||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTTTGCTCACAGCAGCTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||| |||| |||||||| |:: ||||||| :|||:|||:|||||||| ||:|||||||
Db        121 AAAACCTTAAAACAACTAAGTGGGCTCCTTGG-ATAATAGGCTTCTGCCAAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:||| :|||||||||| :|||||:|||| ||||||| ||||||||||||||
Db        180 -CCCAGGTATGGTGAAATAGCCAGACCATTATATACATTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db        239 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||||||:||||||||||||:||||||||:||||||||||     :|||:
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTT---CGTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||| | |||||||||:||:||||||:||||||||||||||||:|||:|||||: |||||
Db        354 GTCAC-GGAAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATAAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||||
Db        413 GCAACCCATGGCATACCTGAGTAAGAAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||||  ||||||||||::||||||||||||| ||||||||||||||||
Db        473 TTTACGGGTAGTGGCAGCAGTAGCAGTTTTAGTATCTGAAGCATTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :||  |||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        533 AAGGTATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:|||||||||||:||||||| ||||||||||:|||||
Db        593 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTGTATTACTTGAAGGGCCA 650


RESULT 368
AB096026
LOCUS       AB096026                543 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-6.
ACCESSION   AB096026
VERSION     AB096026.1  GI:30698502
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hong,K., Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of HERV-W env Fragments in Fetal Brain cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-6"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .543
                     /gene="env"
     repeat_region   1. .543
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             7.0%;  Score 526.7;  DB 1;  Length 543;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  536;  Conservative    4;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
```

```
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTTCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAAC-TCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             |||||||||||||||||||| ||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

Qy       6738 AGA 6740
              | |
Db        541 AAA 543


RESULT 369
AB126108
LOCUS       AB126108                653 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-2.
ACCESSION   AB126108
VERSION     AB126108.1  GI:47169708
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 653)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-2"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .653
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 526.5;  DB 1;  Length 653;
  Best Local Similarity   84.7%;  Pred. No. 0;

  Matches  557;  Conservative   58;  Mismatches   38;  Indels    5;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||||:::|||||||||||||||||   | |||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTCTTACTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|||||||||
Db         121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||  ||||:|||||||||||  :|||||:|||||:|||||||||||||||| |||||||
Db         181 -CCCAGGTGTGGCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTGAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||: |||||||  :|:|||||  :::|||  ||||||||||||||  ||  |
Db         240 CAATACCCATTTAG-AAGATGGGCACCTGAAACAGAAGCGGCTTTCCAGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||: |||||||  |||:|||||||  :|  |  ||||||||:||:| |:
Db         299 GGCCCT--AACCCAAGTTCCAGTGTAAAGCTTGCCAATGGAGGAAGACTTTTCTTTACAT 356

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||  |||||||||:|||:|||||||:|||||||||||||||:|||:||||||:||||||
Db         357 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:||||||||:|||||||||||||||| |||||:|:||||||
Db         416 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 475

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||| ||::||| ||||||||||||||||||||| | ||
Db         476 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGCATCTGAAGCAGTTAAAATAATATAAGG 535

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||| |||||||| ||||| |||||||:||||::|||| || |||||||||| ||||
Db         536 AAGAGATGTTACTGTGCGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 595

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||: || |||||||:||||||||||||||||||:|||||
Db         596 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 653


RESULT 370
AB056499
LOCUS       AB056499                654 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP15-3.
ACCESSION   AB056499
VERSION     AB056499.1  GI:15341227
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP15-3"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .654
                     /gene="pol"
     misc_feature    <1. .>654
                     /gene="pol"

  Query Match             7.0%;  Score 525.5;  DB 1;  Length 654;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   29;  Indels    9;  Gaps    5;

Qy       3092 CAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAA 3151
             | || : :|:|||||:||||||||| | |||| ||||||||||||:||||:|||||||||
Db          2 CCAAACCAAAGCTCAACTCTGCTCAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAA 61

Qy       3152 AGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAA 3211
             ||||||||:|||||||||||||||||:::|||||||| ||||||||||||||||||:|||
Db         62 AGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCCAA 121

Qy       3212 AACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCC 3271
             |||||||||||||||||:|::||| |||||:|| :|||:||||||||||:||||||| |
Db         122 AACCCTAAAGCAACTAAGAGGGTTCTTTGGCATACCAGGCTTCTGCCGAATATGGATT-C 180

Qy       3272 CCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCA 3331
             ||||||: |||:|||||||| |||:|||||:|||||:||||| ||||||||||||||||||
Db         181 CCAGGTACGGCGAAATAGCTAGGCCATTATATACACTAATTTAGGAAACTCAGAAAGCCA 240

Qy       3332 ATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAG 3391
             ||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  | |
Db         241 ATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTGAAGAAGG 300

Qy       3392 GCCTTAAACCCAAGYCCCAGTGT---TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
              |||  ||||||||:||||||||    ||||:||||||||:||||||||||||| :||:|||
Db         301 CCCT--AACCCAAGCCCCAGTGTTAATAAGCTTGCCAACGGGGCAAGACTTTCCTTTATA 358

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
             ::|||||| |||||||||:||:||||||:|||||||||||||||:|||:|||||: |||
Db         359 TGTCACAG-AAAAAACAGAAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACAAGCC 417

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
             |||||||::||||::||||||:||||||||:||||||||||||||||||:|:||||||
Db         418 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 477

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
             |||||: |||||||| ||||||||||::|||||||||||||||||| |||||||||
Db         478 GTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGG 537

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
               | |||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db         538 --AAAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 595

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||:||||||||||||:|||||||||| | ||||||:| |||
Db         596 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTACTGCTTGAAGGACCA 654


RESULT 371
AB115138
LOCUS       AB115138                 541 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE10-4.
ACCESSION   AB115138
VERSION     AB115138.1  GI:41016573
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE10-4"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W env"
     repeat_region   1. .541
                     /note="HERV-W env"

  Query Match             7.0%;  Score 525.4;  DB 1;  Length 541;
  Best Local Similarity   97.8%;  Pred. No. 0;
  Matches  529;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6198 GAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAA 6257
             |||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db          1 GAACAACTTCAGCACAGAAATAAATACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAA 60

Qy       6258 TSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
             | |||| ||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db         61 TTTGGAGATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAC 120

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTC 6377
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         121 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTACCTACCCTC 180

Qy       6378 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 240

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db         241 TATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATA 300

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CAATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 360

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 420

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             |||||||||||||||||||||||||||||||||||||||||||||:|||| ||||||||||
Db         481 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATTGAAGAGCTTT 540

Qy       6738 A 6738
             |
Db         541 A 541


RESULT 372
AB114957
LOCUS       AB114957                 541 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE6-1.
```

```
ACCESSION   AB114957
VERSION     AB114957.1  GI:41016465
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE6-1"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
     repeat_region   1. .541
                     /note="HERV-W env"

  Query Match             7.0%;  Score 525.3;  DB 1;  Length 541;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  534;  Conservative    4;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACT-CCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             |||||||||||||||||||||| ||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

RESULT 373
```

```
AB056489
LOCUS       AB056489                654 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP5-2.
ACCESSION   AB056489
VERSION     AB056489.1  GI:15341217
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP5-2"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .654
                     /gene="pol"
     misc_feature    <1. .>654
                     /gene="pol"

  Query Match             7.0%;  Score 525.2;  DB 1;  Length 654;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   32;  Indels    6;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||  |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||||||:|||||||||||||||||:::|||||||||||||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||| |||||| ||||:|: |||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTGAAGCAATTAAGAGG-TTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: || :|||||||||||:||||:| |||:|||||||||||||||||||||||||
Db        180 -CCCAGGTACGGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:||  |||||:|:||||||:::|||||||||| | |||||  ||  |
Db        239 CAATACCCATTTAGTAGAATGGACACCTGAAGCAGAAGTGGCTTTTCTGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||| | |:|||||||||||:||||||||:|||||||||||||:||:|||:
Db        299 GGCCCT--AACCTATGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATAT 356

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| ||||||||||:||:|||||||:|||||||||||||||| :|||:|| ||: |||||
Db        357 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACATGTCCGAGAGACCAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:| |||||||| ||||||||||:|:||||||
Db        416 GCAACCCGTGGCATACCTGAGTAAGGAAATTTATGTAGTGACAAAGGGTTGGCCTCATTG 475
```

```
Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        476 TTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTAAAGCAGTTAAAATGATACAGGG 535

Qy       3630 -RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
Db        536 AAAGAGATCTTACTGTGTGGCCATCTCATAATGTGAACAGCATACTCACTGCTAAAGGAG 595

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        596 ACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654


RESULT 374
AB115297
LOCUS       AB115297                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP15-1.
ACCESSION   AB115297
VERSION     AB115297.1  GI:41052525
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP15-1"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 525.1;  DB 1;  Length 650;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  557;  Conservative   57;  Mismatches   36;  Indels    9;  Gaps    4;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATC 3148
Db          1 TCCAAACCCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATC 60

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208
Db         61 CAAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCC 120

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
Db        121 CAAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGAT 180

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
Db        181 T-CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAG 234

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
Db        235 CCAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGA 294

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
Db        295 AGGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATA 352

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
Db        353 AGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCT 411

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
Db        412 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATT 471

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
Db        472 GTTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGG 531

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
Db        532 GAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 591

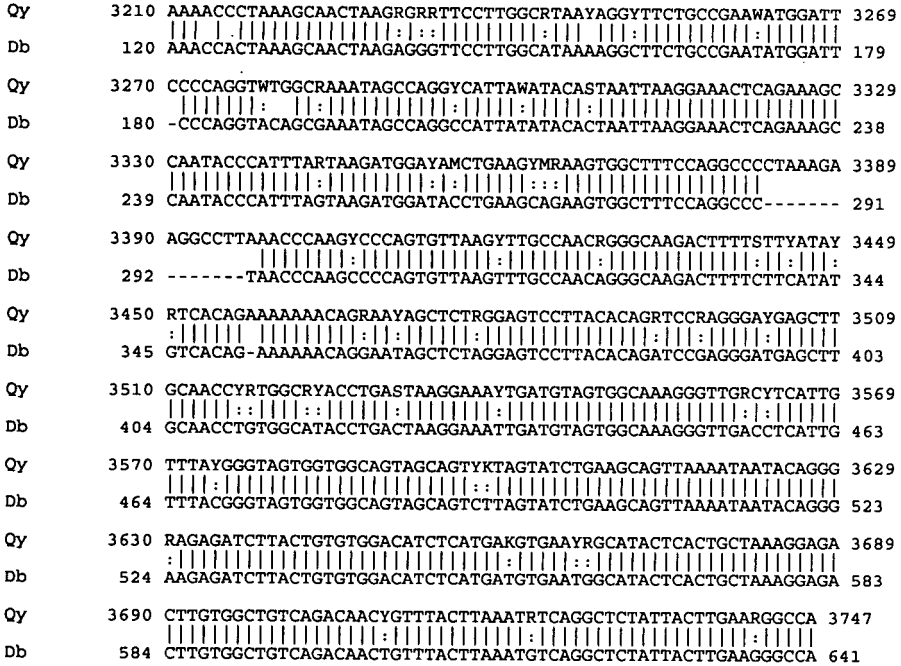
Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        592 ACTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 650


RESULT 375
AB115202
LOCUS       AB115202                641 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-10.
ACCESSION   AB115202
VERSION     AB115202.1  GI:41016632
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 641)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP3-10"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .641
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 524.9;  DB 1;  Length 641;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  556;  Conservative   57;  Mismatches   28;  Indels   17;  Gaps    4;

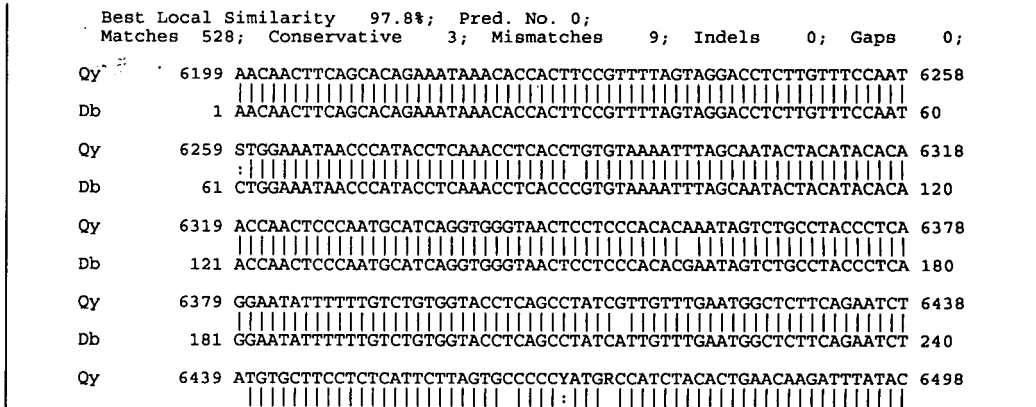
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCCAACCCAAGGGTCAGGTTTGGTCCCAGCAGGTTAAAAGCTTAGGGATAA-ATTATCC 59

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         60 AAAGGCCCCAGGCCTTTCAGTAAGAACCCTATTTAGCCTAAAATGGCTTATCCTCATCCC 119
```



RESULT 376
AB114946
LOCUS AB114946 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE3-16.
ACCESSION AB114946
VERSION AB114946.1 GI:41016454
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE3-16"
/db_xref="taxon:9606"
/tissue_type="testis"
repeat_region 1. .540
/note="HERV-W env"

Query Match 7.0%; Score 524.4; DB 1; Length 540;



RESULT 377
AB115231
LOCUS AB115231 649 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-5.
ACCESSION AB115231
VERSION AB115231.1 GI:41016661
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 649)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .649
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WP10-5"
/db_xref="taxon:9606"
/tissue_type="spleen"
/note="HERV-W pol"

```
     repeat_region   1. .649
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 524.1;  DB 1;  Length 649;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  556;  Conservative   57;  Mismatches   36;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:| |||:|| |||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGTTCAGCTATGCTCAAAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             |||||| |||:||| ||||||||||||:::|||||||||||||||||| ||||||||:|
Db         61 AAAGGCCCCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: |||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||||  |  ||  |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:||||| ||||||:|||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:||||||||||||||||:|||:|||||:||||||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::|||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | |||||||||::|||||||||  |||||||| ||||||||
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||:||||||||||||:|||||||||||||||||| :|||||
Db        592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 649


RESULT 378
AB114975
LOCUS       AB114975                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-7.
ACCESSION   AB114975
VERSION     AB114975.1  GI:41016483
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-7"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             7.0%;  Score 524;  DB 1;  Length 540;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  527;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGTATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:|||||||||||||  ||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAATGAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||  || |||||||||||||||||||||| ||||||||||||||||||||||||
Db        301 AATCATGTTGTACCTAAGCCCCGCAACAAAAGAGTTCCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 379
A46019
LOCUS       A46019                  645 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 8 from Patent EP0674004.
ACCESSION   A46019
VERSION     A46019.1  GI:2300291
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Mandrand,B., Perron,H., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and pathogenic and/or infections agent MSRV2 linked to
            multiple sclerosis, their nucleic components and their applications
  JOURNAL   Patent: EP 0674004-A 8 27-SEP-1995;
            BIO MERIEUX (FR)
COMMENT     Other publication DE 674004T 960919
            Other publication NO 953925 951204
```

```
            Other publication FI 954699 951003
            Other publication AU 1711495 950821
            Other publication CA 2141907 950805
            Other publication FR 2715939 950811
            Other publication FR 2715938 950811
            Other publication FR 2715937 950811
            Other publication FR 2715936 950811.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
Db         121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db         181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db         241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db         301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db         361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db         420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db         477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
Db         535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
Db         594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 380
A57055
LOCUS       A57055                   645 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 8 from Patent WO9628552.
ACCESSION   A57055
VERSION     A57055.1  GI:3713038
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 8 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
Db         121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
Db         181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
Db         241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
Db         301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
Db         361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
Db         420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
Db         477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
Db         535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
Db         594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 381
A60120
LOCUS       A60120                   645 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 8 from Patent WO9706260.
ACCESSION   A60120
VERSION     A60120.1  GI:3715136
KEYWORDS    .
SOURCE      unidentified
```

```
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 8 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||:||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 382
A79470
LOCUS       A79470                   645 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 8 from Patent EP0789077.
ACCESSION   A79470
VERSION     A79470.1  GI:6092478
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 8 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||:||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 383
A80204
LOCUS       A80204                   645 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 8 from Patent WO9521256.
ACCESSION   A80204
VERSION     A80204.1  GI:6092936
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
```

```
            unclassified sequences.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H. and Mallet,F.
  TITLE     MSRV1 VIRUS AND MSRV2 PATHOGENIC AND/OR INFECTIOUS AGENT ASSOCIATED
            WITH MULTIPLE SCLEROSIS, NUCLEIC ACID COMPONENTS AND APPLICATIONS
            OF SAME
  JOURNAL   Patent: WO 9521256-A 8 10-AUG-1995;
            BIO MERIEUX (FR); PERRON HERVE (FR)
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  ||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:|||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 384
A80313
LOCUS       A80313                   645 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 8 from Patent EP0731168.
ACCESSION   A80313
VERSION     A80313.1  GI:6093040
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 8 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  ||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  ||||||||:|||||||| ||:||||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:|||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 385
AR035356
LOCUS       AR035356                 645 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 8 from patent US 5871745.
ACCESSION   AR035356
VERSION     AR035356.1  GI:5952024
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
```

```
  TITLE     Multiple sclerosis related virus
  JOURNAL   Patent: US 5871745-A 8 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  ||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||:::|||||||||||:|||:|||:||||||||||:|  ||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||  |||||:|||||||||:|:||    :::||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:|||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 386
AR036197
LOCUS       AR036197                645 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 8 from patent US 5871996.
ACCESSION   AR036197
VERSION     AR036197.1  GI:5952865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Retrovirus agents MSRV1 and MSRV2 associated with multiple
            sclerosis
  JOURNAL   Patent: US 5871996-A 8 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  ||||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||:::|||||||||||:|||:|||:||||||||||:|  ||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||  :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||  |||||:|||||||||:|:||    :::||||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:|||||||||  ||:||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 387
AR036717
LOCUS       AR036717                645 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 8 from patent US 5800980.
ACCESSION   AR036717
VERSION     AR036717.1  GI:5954573
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Detection of MSRV1 virus and MSRV2 pathogen and/or infective agent
            associated with multiple sclerosis, by nucleic acid hybridization
  JOURNAL   Patent: US 5800980-A 8 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
```

```
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| |||||  |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:|||||||||  ||:|||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:|||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 388
AR076979
LOCUS       AR076979                645 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 8 from patent US 5962217.
ACCESSION   AR076979
VERSION     AR076979.1  GI:10003725
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and MSRV2 pathogen and/or infective agent associated
            with multiple sclerosis, and biopolymer constituents thereof
  JOURNAL   Patent: US 5962217-A 8 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||  |||||||||||||||:| :|||::||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| |||||  |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  ||||||||:|||||||||  ||:|||||||||:||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:|||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||:||||||||:|||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 389
AR094700
LOCUS       AR094700                645 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 8 from patent US 6001987.
ACCESSION   AR094700
VERSION     AR094700.1  GI:10021865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 8 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
```

```
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  |||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 390
AR183935
LOCUS       AR183935                645 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 8 from patent US 6342383.
ACCESSION   AR183935
VERSION     AR183935.1  GI:20227904
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Cell lines and viral isolates associated with multiple sclerosis
  JOURNAL   Patent: US 6342383-A 8 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||:|||||||||||||||:::|||||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  |||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
               | | |||  |||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:||||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::||||||:||||||||:||||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 391
AR343147
LOCUS       AR343147                645 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 8 from patent US 6579526.
ACCESSION   AR343147
VERSION     AR343147.1  GI:33738649
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 8 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             |||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :|||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180
```

```
Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  ||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 392
AR344304
LOCUS       AR344304                645 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 8 from patent US 6582703.
ACCESSION   AR344304
VERSION     AR344304.1  GI:33740245
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 8 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             7.0%;  Score 523.9;  DB 1;  Length 645;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  550;  Conservative   61;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTCAATTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||||||||||||||||:|||  |||||||||||||||:| :|||::|||||||||
Db         61 ACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTTTAGTCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||| :||||||:| |||||||  |:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGAACTCTTAACTTTCCTCACTACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT 3147
             |||||||||:|:|:|||| :|||||||||||| || ||| ||||| |||:||||:|||||
Db        181 TTTCCAAACCAAAGGCTCGGCTCTGCTCACAGGAGATTAGATACTNAGGGCTAAAATTAT 240

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||||||||||||||||:::|||||||||||||||||||||||||||||
Db        241 CCAAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATC 300

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::|||||||||:|||:|||:||||||||||:|  ||
Db        301 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAAACAGA 360

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  :|  ||||||||| :|||||:|||||:|||||| ||||||||||||
Db        361 TT-CCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTANGGAAACTCAGAAA 419

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             ||||||||| |||||:|||||||||:|:||    :::|||||||||||||||||  ||
Db        420 GCCAATACCTATTTAGTAAGATGGACACCT---ACAGAAGTGGCTTTCCAGGCCCTAAAG 476

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||||:|||||||| ||:|||||||:|||||||| ||||:||:||
Db        477 AAGGCCCT--AACCCAAGCCCCAGTGTTCAGCTTGCCAACAGGGCAAGATTTTTCTTTAT 534

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |:: ||||| |||||||||:||:||||||:|||||||||| |||:||  |||||:||||
Db        535 ATGCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGC 593

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
             ||||||||::||| ::|||||||:||||||||:||||||||||||||||||
Db        594 TTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT 645


RESULT 393
AB115216
LOCUS       AB115216                654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-9.
ACCESSION   AB115216
VERSION     AB115216.1  GI:41016646
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-9"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 523.8;  DB 1;  Length 654;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  563;  Conservative   56;  Mismatches   33;  Indels    6;  Gaps    5;
```

```
Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTA-GGRCTAARATTATCC 3149
Db          2 CCAAACAAAAGGGTCAGCTTTGCTCACAGCAGGTTAAATACTTAGGGGGTAAAATTATCC 61

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         62 AAAGGGACCAGGGCCCTCAGTGAGGAACATATCCAGTCTTTACTGGCTTATCCTTATCCC 121

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        122 AAAACCCTAAAGCAATTAAGAG-ATTCCTTGGCATAACAGGTTTCTGCCGAGTATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTATTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGAGTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCACAG-AAAAAACAGGAATAGCTCTAGTAGTCCTTACACAGGTCTGAGGGATGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCACGGCATACCTGAGTAAGGCAATCGATGTAATGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        477 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 536

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        537 AAGAGATCTTACTGTGTGGACTTCTCATGATGTGAACTGCATACTCACTGCTAAAGGAGA 596

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        597 CTTGTGGCTGTCAGACAACCATTTATTTAAATATCAGGCTCTATTACTTGAAGGGCTA 654
```

RESULT 394
AB056488
LOCUS AB056488 654 bp DNA linear PRI 16-JAN-2002
DEFINITION Homo sapiens pol gene, HERV-W family, isolate:HWP5-1.
ACCESSION AB056488
VERSION AB056488.1 GI:15341216
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kim,H.S.
TITLE Sequence and phylogeny of HERV-W pol fragments
JOURNAL AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
PUBMED 11779355
REFERENCE 2 (bases 1 to 654)
AUTHORS Kim,H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .654
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="HWP5-1"
/db_xref="taxon:9606"
/chromosome="5"
/note="HERV-W pol
human endogenous retrovirus family :HERV-W"
gene 1. .654
/gene="pol"
misc_feature <1. .>654
/gene="pol"

Query Match 7.0%; Score 523.6; DB 1; Length 654;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 561; Conservative 59; Mismatches 33; Indels 6; Gaps 5;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTTACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTGAAGCAATTAAGAGG-TTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        180 -CCCAGGTACGGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        239 CAATACCCATTTAGTAGAATGGACACCTGAAGCAGAAGTGGCTTTTCTGGCCCTAAAGAA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        299 GGCCCT--AACCTATGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATAT 356

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        357 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACATGTCCGAGAGACCAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        416 GCAACCCGTGGCATACCTGAGTAAGGAAATTTATGTAGTGACAAAGGGTTGGCCTCATTG 475

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        476 TTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTAAAGCAGTTAAAATGATACAGGG 535

Qy       3630 -RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
Db        536 AAAGAGATCTTACTGTGTGGCCATCTCATAATGTGAACAGCATACTCACTGCTAAAGGAG 595

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        596 ACTTGTGGCTGTCAAACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 654
```

RESULT 395
AB115299
LOCUS AB115299 655 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP15-4.
ACCESSION AB115299
VERSION AB115299.1 GI:41052527
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.

```
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 655)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP15-4"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W pol"
     repeat_region   1. .655
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 523.5;  DB 1;  Length 655;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  554;  Conservative   60;  Mismatches   40;  Indels    5;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
                 |||| ||:|:|:| |||:|| |||||||||||||||||||||||||:||||:|||||||
Db             1 TCCAGACCAAAGGTTCAGCTTTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
                 | |  | |||:||||||||||||||:::|||||||||||||||||||||||| |||:|
Db            61 AGACACCCCAGGGCCCTCAGTGAGGAATGCATCCAGCCTATACTGGCTTATCCTTATCCC 120

Qy       3210 -AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
                  |||||| |||||||| ||||:|::|||||||||:|||:|||:||||||||||:||||||
Db            121 AAAAACCATAAAGCAATTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGAT 180

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
                 | |||||||:  | :||||||||||||:||||||:| |||:|||||||||||||||||||||
Db            181 T-CCCAGGTACAGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAG 239

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
                 ||||||||||||:||||| |||:|:|||||:::|||||||||||||||||||  ||
Db            240 TGAATACCCATTTAGTAAGAAGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGA 299

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
                 | | |||  || |||||:|| |||||||||:|||||| :|||||||||||||:||:||
Db            300 AGGCCCT--AATCCAAGCCCTAGTGTTAAGCTTGCCAAAGGGGCAAGACTTTTCTTTATG 357

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
                 ::||||||| |||||||:||:||||||:||| ||||||||||:|||:|||||: || |
Db            358 TGTCACAGA-AAAAACAGGAATAGCTCTAGGATTCCTTACACAGGTCCGAGGGACCAGGT 416

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
                 |||||||::||||::||||||:|||||||:|||||||||||||||||||:|:|||||
Db            417 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 476

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
                 |||||:|||||  |||  |||||||||||::|||||||||||||||||||||| |||||||
Db            477 GTTTATGGGTAATGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGG 536

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
                 |:|||||||||||||||||  |||||||||:||| |::|||||||||||||||||||||
Db            537 GAAGAGATCTTACTGTGTGGATGTCTCATGATGTGGACGGCATACTCACTGCTAAAGGAG 596

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                 ||||  |||||||||||||:|||| ||||||:|||||||||||||||||:|||||
Db            597 ACTTACGGCTGTCAGACAACCGTTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 655


RESULT 396
AB115271
```

```
LOCUS       AB115271                657 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP6-5.
ACCESSION   AB115271
VERSION     AB115271.1  GI:41052499
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 657)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP6-5"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W pol"
     repeat_region   1. .657
                     /note="HERV-W pol"

  Query Match             7.0%;  Score 522.1;  DB 1;  Length 657;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  567;  Conservative   59;  Mismatches   28;  Indels    7;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
                 |||||||:|:|:|||||:||||||||||||||||||||||||||||:||||:|||||||
Db             1 TCCAAACGAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
                 ||||||||||:|||||||||||||||:::|||||||||||||||||||| |||||:|
Db            61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATTCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
                 .  |||||||||||||||||||:| :||||||||:|||:|||:| |||| |||:|||||||
Db           121 AAAACCCTAAAGCAACTAAGAGCATTCCTTGGCATAACAGGCTGCTGCTGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                 |||||||:  | :||||||||||||:|||||:| |||:|||||||||||||||||||||
Db           181 -CCCAGGTACAGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAGC 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                 |||||||||||||:|||||||:|:|| |||:::||| ||||||||||||   || |
Db           240 CAATACCCATTTAGTAAGATGGACACCTTAAGCAGAAGCGGCTTTCCAGGCCTTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                 | |||  |||||||:|||||| ||||:|||||||:||||||||||||||||||:||:|
Db           300 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAACAGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
                 :|||||||  ||||||:||:|||||:|||||||||||||||:|| :|||||:||||||
Db           358 GTCACAGAAG-AAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCTGAGGGATGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
                 ||||||::||||::||||||:||||||||:|||||||||||||||||||||:|:||||||
Db           417 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAG-TTAAAATAATACA-G 3627
                 ||||:||||||||||  ||||||||||::|||||||||||| || |||||||||||| |
Db           477 TTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGTAGTTTAAAATAATACAGG 536
```

```
Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:|||||||||||||||| |||||||||||:|||||::||||| ||||||||||||||
Db        537 GGAAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATAGTCACTGCTAAAGGA 596

Qy       3688 GACTTGTGGCTGTC-AGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
               ||||| ||||||| ||||||:||||||||||||: ||||||||||||||||:||||
Db        597 AACTTGGGGCTGTCAAGACAACTGTTTACTTAAATACCAGGCTCTATTACTTGAAGGGCC 656

Qy       3747 A 3747
              |
Db        657 A 657


RESULT 397
AB079503
LOCUS       AB079503                541 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-3.
ACCESSION   AB079503
VERSION     AB079503.1  GI:22122077
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-3"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .541
                     /rpt_family="HERV-W"

  Query Match             7.0%;  Score 521.5;  DB 1;  Length 541;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  532;  Conservative    4;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 AT-GTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
              || |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATA 300

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        301 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGGGTACCCATTCTTCCTTTTGTTATAGG 360

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
              |||||||||||| |||| ||||||||||||||||||||| ||||||||||||||||||||
Db        361 AGCAGGAGTGCTGGGTGTACTAGGTACTGGCATTGGCGGCATCACAACCTCTACTCAGTT 420

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        481 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

Qy       6738 A 6738
              |
Db        541 A 541


RESULT 398
AB115099
LOCUS       AB115099                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE2-1.
ACCESSION   AB115099
VERSION     AB115099.1  GI:41016534
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE2-1"
                     /db_xref="taxon:9606"
                     /cell_line="PFSK-1"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             7.0%;  Score 520.7;  DB 1;  Length 537;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  531;  Conservative    4;  Mismatches    2;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
```

```
                                                
```

```
                ||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
                |||||||||||||||||| :|:: |||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                |||||||: |||:|||||||||| |||||:|     ||||||||||||||||||||
Db         181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                |||||||||||||:|||||||||:|:|||| |:::||| ||||||||||| |  ||  |
Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                | |||  ||||||||:|||||||||||||:||||| | :|||||||||||||:||:|||:
Db         295 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
                :||| | |||||| |||:||:||||||:||||||||||||||||:|||:||||||:||||||
Db         353 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
                ||||||::||||::|||||| ||||||||:||||||||||||||||||||:|:||||||
Db         412 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
                ||||:|||||||||  ||||||||||::|||||||||||||||| ||||||||||||
Db         472 TTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
                :|||||| ||||||||||||||||||||:|||||::|||||| ||||||||||||||
Db         532 AAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGT-CAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                |||||||||| || |||||:||||||||||:||||||||||||||||||:|||||
Db         592 CTTGTGGCTGTCCAAACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 650


RESULT 401
BV485214/c
LOCUS       BV485214                654 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S215P61296FD7.T0 Clara Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV485214
VERSION     BV485214.1  GI:62328884
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 654)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 654
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Clara"
     STS             <1. .>654

  Query Match             6.9%;  Score 520;  DB 1;  Length 654;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  557;  Conservative   33;  Mismatches   60;  Indels    4;  Gaps    2;

Qy       3872 TAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTA 3931
                ||||:|||||:||||||||||:|| |||||||||||| ||||||||||||  ||| ||
Db         654 TAGAGGTTCCCTTGACTGATCCCGA-CTTCAACTTGTACACTGATGGAAGTTTATTTTTA 596

Qy       3932 GAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGT 3991
                ||||||||||||  |||||:|   |||||||||||||||||||||||||:||||||||
Db         595 GAAAAAGGACTTTAAAAAGCGTATTATGCAGTGGTCAGTGATAATGGAATACTTGAAAGT 536

Qy       3992 AATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCA 4051
                |||||||||||||||||||||||||||:||||:|||||| ||||||||:|||:|  ||
Db         535 AATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAAATAATAGCCCTCACTCAGGTG 476

Qy       4052 CTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTA 4111
                |||||||||||||| : |||||||:|||||||||||||| ||||:|:|| |||:|||||
Db         475 CTAGAATTAGGAGAAAGGAAAAGGGTAAATATATATACAGTCTCTAAATACGCTTACCTA 416

Qy       4112 GTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACA 4171
                ||| ||||||||||::||||||||:|:|||||||||||||||||:|| :||  |
Db         415 GTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAAGGACAA 356

Qy       4172 CCTATCAMACATCAGGAAGCCATTAGGARATTATTAYTGGCWGTACAGAAACCTARAGAG 4231
                |||||||:|||||||||||||||||||:||||||:||||:| |||||| |||:||||
Db         355 CCTATCAAACATCAGGAAGCCATTAGGAGATTATTATTGGCTGCACAGAAATCTAAAGAG 296

Qy       4232 GTGGMAGTCTTACACTGCYGGGGTCATCANAAAGGAAAGRAAAGGGAAATASAAGRGAAY 4291
                ||||:| |||||||| |:|||||||| |||||||||:||||||||||:|||:|||:
Db         295 GTGGCAATCTTACACTACTGGGGTCATCAGAAAGGAAAGGAAAGGGAAATAGAAGGGAAC 236

Qy       4292 TGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCTCCATTAGAAATGCTT 4351
                ||||||  :|||||||||:|||||||  |||||||||||||||| ||||||||||
```

```
Db        235 CGCCAAGTGGATATTGAAGCCAAAAGAGCCACAAGGCAGGACCCTCCATAAGAAATGCTT 176

Qy       4352 ATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCA 4411

Db        175 ATATAAGGACCCCTAGTATGGGGTAATCCCCTCCAGGAACCCAAGCCCCAGTTCTCAGCA 116

Qy       4412 GGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTTAGCC 4468

Db        115 AGAGAGAGAGAATGGGGAGCCTCACAAGGACATAGTTTCCTCCCCTCAGGATGGCTAGCC 56

Qy       4469 ACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAA 4522

Db         55 ACTGAAGAAGGGAAAATAATTTTGTCTGCAGCTAACCCATGGAAATTATTTAAA 2


RESULT 402
AB079504
LOCUS       AB079504                541 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-4.
ACCESSION   AB079504
VERSION     AB079504.1  GI:22122078
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-4"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .541
                     /rpt_family="HERV-W"

  Query Match             6.9%;  Score 519.9;  DB 1;  Length 541;
  Best Local Similarity   98.2%;  Pred. No. 0;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTC-TTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497

Db        241 ATGTGCTTCCTCTCATTCTTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATA 300

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557

Db         301 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 360

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617

Db         361 AGCAGGAGTGCTGGGTGTACTAGGTACTGGCATTGGCGGCATCACAACCTCTACTCAGTT 420

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677

Db         421 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTTCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737

Db         481 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

Qy       6738 A 6738

Db         541 A 541


RESULT 403
AB115294
LOCUS       AB115294                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP11-4.
ACCESSION   AB115294
VERSION     AB115294.1  GI:41052522
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP11-4"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 519.9;  DB 1;  Length 650;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  553;  Conservative   57;  Mismatches   38;  Indels    9;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149

Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209

Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269

Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180
```

```
Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        592 CTTGGGGCTGTCAAACACCTGTTTACTTAAATATCAGGCTCTATTACTTTGAAGGTC 648
```

RESULT 404
AB056491
LOCUS AB056491 650 bp DNA linear PRI 16-JAN-2002
DEFINITION Homo sapiens pol gene, HERV-W family, isolate:HWP6-4.
ACCESSION AB056491
VERSION AB056491.1 GI:15341219
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kim,H.S.
TITLE Sequence and phylogeny of HERV-W pol fragments
JOURNAL AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
PUBMED 11779355
REFERENCE 2 (bases 1 to 650)
AUTHORS Kim,H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259, Fax:82515812962)

| FEATURES | Location/Qualifiers |
|---|---|
| source | 1. .650 |
| | /organism="Homo sapiens" |
| | /mol_type="genomic DNA" |
| | /isolate="HWP6-4" |
| | /db_xref="taxon:9606" |
| | /chromosome="6" |
| | /note="HERV-W pol |
| | human endogenous retrovirus family :HERV-W" |
| gene | 1. .650 |
| | /gene="pol" |
| misc_feature | <1. .>650 |
| | /gene="pol" |

Query Match 6.9%; Score 519.6; DB 1; Length 650;
Best Local Similarity 84.2%; Pred. No. 0;

Matches 554; Conservative 56; Mismatches 40; Indels 8; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACAAAATGCTCAGCACTGCTCACAGAGTGGTTAAAACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACAAGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACAGCAAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCAGAAAGC 235

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        236 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCAGCTTTTGAGGTCTTAAAGAA 295

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        296 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGGGGCAAGACTTTTCCTTATAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        354 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAGGGACGAGCTT 412

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        413 GCAACCCGTGGCATACCTGAGAAAGGAAAGTGATGTGGTGGCAAAGGGTTGGCCTCATTG 472

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        473 TTTACAGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 532

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        533 AAGAGATCTTACTGTGTGGACATCTCATGATGTAAATCACATACTCACTGATAAAGGAGA 592

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        593 CTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 650
```

RESULT 405
AB054085
LOCUS AB054085 537 bp DNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens DNA, human endogenous retrovirus HERV-W family sequence, isolate:HepG2-1.
ACCESSION AB054085
VERSION AB054085.1 GI:14780050
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
TITLE Molecular cloning and phylogenetic analysis of new human endogenous retrovirus HERV-W family in cancer cells
JOURNAL Curr. Microbiol. (2001) In press
REFERENCE 2 (bases 1 to 537)
AUTHORS Lee,J. and Kim,H.-S.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2001) Joo-Mi Lee, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers

```
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HepG2-1"
                     /db_xref="taxon:9606"
                     /cell_line="HepG2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.9%;  Score 519.1;  DB 1;  Length 537;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  530;  Conservative    4;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       298 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 406
AB114963
LOCUS       AB114963                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-4.
ACCESSION   AB114963
VERSION     AB114963.1  GI:41016471
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES            Location/Qualifiers
     source         1. .537
                    /organism="Homo sapiens"
                    /mol_type="other RNA"
                    /isolate="WE7-4"
                    /db_xref="taxon:9606"
                    /tissue_type="lung"
     repeat_region  1. .537
                    /note="HERV-W env"

  Query Match             6.9%;  Score 519.1;  DB 1;  Length 537;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  530;  Conservative    4;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       118 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db       238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       298 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 407
BD136195
LOCUS       BD136195                635 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION   BD136195
VERSION     BD136195.1  GI:23231140
KEYWORDS    JP 2002509437-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
            Perron,H. and Mandrand,B.
  TITLE     Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: JP 2002509437-A 5 26-MAR-2002;
```

```
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002509437-A/5
            PD   26-MAR-2002
            PF   07-JUL-1998 JP 1999508255
            PR   07-JUL-1997 FR     97/08816
            PI   GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC  PI
              BEDIN,
            PI   MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,
            PI   BERNARD MANDRAND
            PC   C12N15/48,C12Q1/70,C07K14/15,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Retroviral nucleic material and nucleotide fragments, in  CC
                 particular,
            CC   associated with multiple sclerosis and/or
            rheumatoid arthritis,
            CC              for
            CC   diagnostic, prophylactic and therapeutic uses FH   Key
                Location/Qualifiers
            FT   source          1. .635
            FT                   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .635
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
Db            61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
Db           121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
Db           181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
Db           241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
Db           301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
Db           361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
Db           421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
Db           481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db           541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
Db           601 ATTAAATCTTGCAACTG 617


RESULT 408
CS248638
LOCUS       CS248638                635 bp    DNA     linear   PAT 09-JAN-2006
DEFINITION  Sequence 6 from Patent EP1612270.
ACCESSION   CS248638
VERSION     CS248638.1  GI:84662034
KEYWORDS    .                 .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Parahnos-Baccala,G.
  TITLE     Retroviral nucleic material and nucleotide fragments, in particular
            associated with multiple sclerosis and/or rheumatoid arthritis, for
            diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: EP 1612270-A 6 04-JAN-2006;
            BIOMERIEUX (FR)
FEATURES              Location/Qualifiers
     source           1. .635
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
Db            61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
Db           121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
Db           181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
Db           241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
Db           301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
Db           361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
Db           421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
Db           481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db           541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600
```

```
Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 409
AR344385
LOCUS       AR344385                635 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 102 from patent US 6582703.
ACCESSION   AR344385
VERSION     AR344385.1  GI:33740326
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 102 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
              |||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:|||||| | || |||||||||||||  |||||||||||||  |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              ||||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              ||||||||||||||||||||||   |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||||| |||||| ||||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | ||||||||||||||||||||||||:|||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||||| ||||||||||  ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
Db        601 ATTAAATCTTGCAACTG 617


RESULT 410
AX001024
LOCUS       AX001024                635 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 6 from Patent WO9902666.
ACCESSION   AX001024
VERSION     AX001024.1  GI:7241262
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Ott,C. and Bedin,F.
  TITLE     RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR
            ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR
            DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
  JOURNAL   Patent: WO 9902666-A 6 21-JAN-1999;
            BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             6.9%;  Score 517.4;  DB 1;  Length 635;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  552;  Conservative   11;  Mismatches   42;  Indels   12;  Gaps    1;

Qy       6976 CCCTGTATCTTTRACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA 7035
              |||||||||||:|||||||||||||| |||||||||||||||||| || ||||:||||||
Db          1 CCCTGTATCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA 60

Qy       7036 ------------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 7083
                          |||||||||| | ||||| |||||| |||||||||| ||||  |||||
Db         61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Qy       7084 CTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAG 7143
              |||||||||||||:|||||| | || |||||||||||||  |||||||||||||  |||
Db        121 CTGGACCGGCCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG 180

Qy       7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
              ||||||||| |||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db        181 GAAATCTCAACTGCACAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Qy       7204 SGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
              : || |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        241 CATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              ||||||||||||||||||||||   |: ||||||||:|||||||||:||||||||||||
Db        301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||||| |||||| ||||||||||||||||:||||||| ||||||||||||||||||
Db        361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| | ||||||||||||||||||||||||:|||||:||
Db        421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 480

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| |||||||||| |||||||||||||||||| |  ||||||||||||||||
Db        481 AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 540

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||||||||||||||||| ||||||||||  ||||||||||||
Db        541 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTTTCACTCTATTTCACTCT 600

Qy       7564 ATTAAATCTTGCARCTG 7580
              |||||||||||||:|||
```

```
Db        601 ATTAAATCTTGCAACTG 617


RESULT 411
BD196265
LOCUS       BD196265                591 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196265
VERSION     BD196265.1  GI:33006035
KEYWORDS    JP 2002512530-A/31.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 31 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/31
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC            diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .591
            FT                   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .591
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:||||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              |||||||||||||||||||||||||||  ||||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||:|||||||||||||||||:|||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |

Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              || ||||||||||||||||:|||||||||||| |||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              |||||·|||||||| |||||||||||||||||||||||||:|||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 412
BD267476
LOCUS       BD267476                591 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267476
VERSION     BD267476.1  GI:33077244
KEYWORDS    JP 2002534980-A/17.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 17 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/17
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .591
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .591
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:||||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              |||||||||||||||||||||||||||  ||||||||||||||||   ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              ||||||||||||||:||||||||||||||||||||:|||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
                || ||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| ||||||||||||||||||||||||||:||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 413
AX000986
LOCUS       AX000986                 591 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 31 from Patent WO9902696.
ACCESSION   AX000986
VERSION     AX000986.1  GI:7241228
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 31 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              ||||||||||||||||||||||||||||||||  ||||||||||||||    ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||||:|||||||||||||||||||:|||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
                || ||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| ||||||||||||||||||||||||||:||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 414
AX027469
LOCUS       AX027469                 591 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 19 from Patent FR2788784.
ACCESSION   AX027469
VERSION     AX027469.1  GI:10188433
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 19 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              ||||||||||||||||||||||||||||||||  ||||||||||||||    ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||||:|||||||||||||||||||:|||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
```

```
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              || |||||||||||||||||:|||||||||||||| |||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| |||||||||||||||||||||||||:|||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 415
AF072498
LOCUS       AF072498                591 bp    DNA     linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens multiple sclerosis associated retrovirus env probe
            Penv-C15 sequence.
ACCESSION   AF072498
VERSION     AF072498.1  GI:4262282
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 591)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="isolated from synoviocyte cultures of a rheumatoid
                     arthritis (RA) patient"
     misc_feature    1. .591
                     /note="MSRV env probe Penv-C15"
     repeat_region   1. .591
                     /rpt_family="multiple sclerosis associated retrovirus"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 591;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  530;  Conservative   12;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6467 CYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACA 6526
              |:|||:|||||||||||||||||||||||||| | ||||| || ||||||||| |||||
Db          1 CCATGGCCATCTACACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCCACAACA 60

Qy       6527 AAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTG 6586
              |||||||||||||||||||||||||||  ||||||||||||||    ||||||||||
Db         61 AAAGAGTACCCATTCTTCCTTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTG 120

Qy       6587 GCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATG 6646
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 GCATTGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATG 180

Qy       6647 GGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 6706
              | |||||||||| |||| | ||||||||||||||||||||||||||||||||||||||
Db        181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAG 240

Qy       6707 CAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGACTYGCTAACCGCTGARAGAGGGGGAA 6766
              |||||||||||||||:||||||||||||||||||:||||||||  |:||||||||||
Db        241 CAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAAGAGGGGGAA 300

Qy       6767 CCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTG 6826
              |||||||||||||||| |||||| |||||||||||||||||||| |||| |||||||
Db        301 CCTGTTTATTTTTAGGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAATTGTCACTG 360

Qy       6827 AGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACACTG 6886
              ||||||||:|||||||||:|||||||||||||| |||:||||||:||||||| |||||| |
Db        361 AGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCG 420

Qy       6887 GACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCCTTCTTAGGACCTCTAG 6946
              || |||||||||||||||||:|||||||||||||| |||||||||||||||||||||||||
Db        421 AACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCCTTCTTAGGACCTCTAG 480

Qy       6947 CAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTGTTAACTTTG 7006
              ||||| |||||||| |||||||||||||||||||||||||:|||||||||||||| ||||
Db        481 CAGCTCTAATATTGTTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAAGTTTG 540

Qy       7007 TCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||||||||||||| ||||||||:|| ||||| ||||
Db        541 TCTCTTCCAGAATTGAAGCTGTAAAGCTACAGATGG 576


RESULT 416
AB115292
LOCUS       AB115292                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP11-1.
ACCESSION   AB115292
VERSION     AB115292.1  GI:41052520
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP11-1"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 516.8;  DB 1;  Length 650;
  Best Local Similarity   84.7%;  Pred. No. 0;
  Matches  558;  Conservative   57;  Mismatches   34;  Indels   10;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:||| ||||||||||||:::|||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
```

```
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329

Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTCAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689

Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACY-GTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        592 CTTGTGGCTGTCAGACAACTGGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 650
```

```
RESULT 417
AB114925
LOCUS       AB114925                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-14.
ACCESSION   AB114925
VERSION     AB114925.1  GI:41016433
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-14"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.9%;  Score 516;  DB 1;  Length 540;
  Best Local Similarity   96.7%;  Pred. No. 0;
```

```
  Matches  522;  Conservative    4;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 418
AB115266
LOCUS       AB115266                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP5-4.
ACCESSION   AB115266
VERSION     AB115266.1  GI:41052494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP5-4"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W pol"
     repeat_region   1. .650
```

/note="HERV-W pol"

Query Match 6.9%; Score 515.8; DB 1; Length 650;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 557; Conservative 57; Mismatches 34; Indels 10; Gaps 5;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTG-TGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        592 CTTGTTGGCTGTCAAACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAAGTCC 649
```

RESULT 419
AB115201
LOCUS AB115201 649 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-9.
ACCESSION AB115201
VERSION AB115201.1 GI:41016631
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 649)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .649
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WP3-9"
/db_xref="taxon:9606"
/tissue_type="testis"
/note="HERV-W pol"
repeat_region 1. .649
/note="HERV-W pol"

Query Match 6.9%; Score 515.7; DB 1; Length 649;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 550; Conservative 58; Mismatches 41; Indels 9; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAGGGTCAGCTTTGGTCAAAGCAGGTTAAAAAGGAAGGGCCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCCCCAGGGCAATCAGTGAGGAACGTATCCAGCCTAAACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAACCCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGCGAAATAGCCAGACCATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        592 CTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649
```

RESULT 420
AB115270
LOCUS AB115270 652 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP6-4.
ACCESSION AB115270
VERSION AB115270.1 GI:41052498
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 652)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP6-4"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W pol"
     repeat_region   1. .652
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 515.3;  DB 1;  Length 652;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  556;  Conservative   59;  Mismatches   36;  Indels    9;  Gaps    5;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCT-AARATTAT 3147
             | | || | |:|:|||||:||||||||| ||||||||||||| ||||:   ||:|||||
Db             1 TCCAAACCAAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGGCAAAATTAT 60

Qy       3148 CCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
             ||||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||
Db            61 CCAAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATC 120

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
             :|||||||||||||||||||||:|::|||||||||:|||:| |:|||||||||:|||||
Db            121 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATATGGA 180

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
             || |||||||:  ||:|||||| ||||:|||||:| |||:    |||||||||||||||||
Db            181 TT-CCCAGGTACAGCGAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCAGAAA 235

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
             |||||||||||||||:|||||||||:|:||||||:::|||||||||| |||||||  ||
Db            236 GCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAG 295

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
              | | |||  |||||||:|||||||| ||||:||||||| :|||||||||||||:||:||
Db            296 AAGGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTAT 353

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
             |::|| || ||| |||||||:||:|||||||:|||||||||| ||||:|||:|||||: |||
Db            354 ATGTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGC 412

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
             ||||||||::||||::||||||:||||||||:||||||| |||||||||||||:|:||||
Db            413 TTGCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCAT 472

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
             ||||||:|||||||||   ||||||||||::|||||||||||||||||||| ||||||||
Db            473 TGTTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAG 532

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:|||||||||||||||| ||||||||||:|||||::|||||||||||||||||||||||
Db            533 GGAAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGA 592

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             || |||||||||||||||||:||||||||||:|||||||||||||||||| :|||||
Db            593 GATTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 652
```

```
RESULT 421
AB115207
LOCUS       AB115207                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP5-9.
ACCESSION   AB115207
VERSION     AB115207.1  GI:41016637
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP5-9"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 514.9;  DB 1;  Length 685;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  565;  Conservative   59;  Mismatches   30;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
             |||||||:|:|:|||||:|||||||||||||||||              ||||||||||||
Db             1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
             :||||:||||||||||||||:|||||||||||||||||:::|||||||||||||||||||
Db            61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
             ||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db           121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAGC 180

Qy       3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                         |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db           181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db           240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             |||:::| |  |||||||||||||   ||  |  |||  ||||||||:||||||||||||
Db           300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :|||||| |:|||||||||||:||: ||::|||||| |||||| ||:||:|||||||:|
Db           358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||||||:|||:|||||: ||||||||||::||||::|||||||:||||||||
Db           417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476
```

```
Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
              :|||||||||||||||||||||:|:||||||||||:|||||||| ||||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
              |||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
              |:|||||: |||||||||||||||||||||||||||||||||||||||:|||| ||||
Db        597 ATGTGAACCGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||:||||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 422
AB114953
LOCUS       AB114953                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-5.
ACCESSION   AB114953
VERSION     AB114953.1  GI:41016461
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-5"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.9%;  Score 514.4;  DB 1;  Length 540;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  521;  Conservative    4;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db         241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || ||||||||| ||||||||||||||||||||||||||||||||  ||
Db         301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db         421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db         481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 423
AB115156
LOCUS       AB115156                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE14-2.
ACCESSION   AB115156
VERSION     AB115156.1  GI:41016591
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE14-2"
                     /db_xref="taxon:9606"
                     /cell_line="LOX-IMVI"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.9%;  Score 514.4;  DB 1;  Length 540;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  521;  Conservative    4;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180
```

```
Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        421 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||| ||||||||||||||||||:||||| ||||| ||||
Db        481 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 540


RESULT 424
AB115269
LOCUS       AB115269                650 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP5-10.
ACCESSION   AB115269
VERSION     AB115269.1  GI:41052497
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP5-10"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W pol"
     repeat_region   1. .650
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 514.2;  DB 1;  Length 650;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  556;  Conservative   57;  Mismatches   35;  Indels   10;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| |||||||||||:::|||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||| :|:: |||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||||| |  ||  |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||||||:|||||| ||||||:|||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:|||||||||||||||:|||:|||||:||||||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | ||||||||||::|||||||||| |||||||| ||||||| |||||||
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATR-TCAGGCTCTATTACTTGAARGGCC 3746
             |||| |||||||||||||:|||||||||||: |||||||||||||||| |:|| |
Db        592 CTTGGGGCTGTCAGACAACTGTTTACTTAAATATTCAGGCTCTATTACTTGGAAGGTC 649


RESULT 425
AB115295
LOCUS       AB115295                651 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP11-5.
ACCESSION   AB115295
VERSION     AB115295.1  GI:41052523
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 651)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP11-5"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W pol"
     repeat_region   1. .651
                     /note="HERV-W pol"
```

```
  Query Match             6.9%;  Score 514.2;  DB 1;  Length 651;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches  556;  Conservative   57;  Mismatches   35;  Indels   10;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| ||||||||||||:::|||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||||||||||| :|:: |||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||| |  ||  |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  |||||||:|||||  ||||||:|||||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:|||||||||||||||||:|||:|||||:||||||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | ||||||||||::|||||||||  |||||||| ||||||||
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTTGTGGCTGTCAGACAACYGTTT-ACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             ||||||||||||||||||:|||| ||||||||:|||||||||||||  |:|| |
Db        592 CTTGTGGCTGTCAGACAACTGTTTAACTTAAATATCAGGCTCTATTACTTTGAAGGTC 649


RESULT 426
AB126118
LOCUS       AB126118                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-6.
ACCESSION   AB126118
VERSION     AB126118.1  GI:47169718
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-6"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 513.2;  DB 1;  Length 638;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  547;  Conservative   59;  Mismatches   32;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:||||:||||| ||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||  |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||: ||||||||||:|||||:|||||:||||||||||||||||||||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA--------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||||| :|||||| |:|||||  |||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:|||||||:|||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:|||||||||:|||||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||| ||::|||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||| |||||||:|||||::|||| || ||||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||: || ||||||||:|||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 427
AB126122
LOCUS       AB126122                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP7-1.
ACCESSION   AB126122
VERSION     AB126122.1  GI:47169722
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
```

```
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP7-1"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 513.2;  DB 1;  Length 638;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  547;  Conservative   59;  Mismatches   32;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||| |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||: |||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||:|||||||||:|:|||||:::|||||||||||||||||||  ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||| :|||||| |:|||| |||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:|||||||:|||||||||||||||:|| :|||||:|||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||| ||::|||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db         521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||: || |||||||:|||||||||||||||||:|||||
Db         581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 428
AB126127
```

```
LOCUS       AB126127                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP8-4.
ACCESSION   AB126127
VERSION     AB126127.1  GI:47169727
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP8-4"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 513.2;  DB 1;  Length 638;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  547;  Conservative   59;  Mismatches   32;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             ||||||||||| |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||: |||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||:|||||||||:|:|||||:::|||||||||||||||||||  ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||| :|||||| |:|||| |||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:|||||||:|||||||||||||||:|| :|||||:|||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||| ||::|||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520
```

```
Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:||||::|||| || |||||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:|||||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 429
AB126130
LOCUS       AB126130                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP9-2.
ACCESSION   AB126130
VERSION     AB126130.1  GI:47169730
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP9-2"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.9%;  Score 513.2;  DB 1;  Length 638;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches  547;  Conservative   59;  Mismatches   32;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| ||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||| |||||||:|:: |||||||:|||:|||:||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: |||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||  :|||||  |:|||||  |||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||||:|| :|||||:||||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              |||||||::|||| :|||| |:|||||||:|||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:|||||::|||| || |||||||||| ||||
Db         521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:|||||||||||||||||||:|||||
Db         581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 430
AB114934
LOCUS       AB114934                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-6.
ACCESSION   AB114934
VERSION     AB114934.1  GI:41016442
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-6"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.8%;  Score 512.9;  DB 1;  Length 539;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  527;  Conservative    3;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||
Db         121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
```

```
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        240 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539


RESULT 431
AB054087
LOCUS       AB054087                537 bp    DNA     linear   PRI 06-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:JUR-3.
ACCESSION   AB054087
VERSION     AB054087.1  GI:15005017
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
  TITLE     Molecular cloning and phylogenetic analysis of new human endogenous
            retrovirus HERV-W family in cancer cells
  JOURNAL   Curr. Microbiol. (2001) In press
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Lee,J. and Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="JUR-3"
                     /db_xref="taxon:9606"
                     /cell_line="Jurkat"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.8%;  Score 512.7;  DB 1;  Length 537;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches  526;  Conservative    4;  Mismatches    7;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
```

```
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 432
AB115206
LOCUS       AB115206                651 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP5-8.
ACCESSION   AB115206
VERSION     AB115206.1  GI:41016636
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 651)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP5-8"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
                     /note="HERV-W pol"
     repeat_region   1. .651
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 512.1;  DB 1;  Length 651;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  554;  Conservative   58;  Mismatches   37;  Indels    9;  Gaps    5;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCA 3150
              |||| | |:| | |||:|| ||||| |||||||||||| | |||||: |||:||||||||
Db          2 CCAACCAAAATGGTCAGCTTTGCTCCCAGCAGGTTAAAAATTTAGGGGTAAAATTATCCA 61

Qy       3151 AAGGCACCA-RGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||| | | :|||||||||||||||:::|||||||||||||||||||||||||||:|
```

```
Db        62 AAGGCCCAAGGGGCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCC 121

Qy      3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269

Db       122 AAACCCCTAAAGCAACTAAGAGGGTTCCTTGGCATACCAAGCTTCTGCCGAATATGGATT 181

Qy      3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329

Db       182 -CCCAGGTACAGCGAAATAGTCAGGCCATTATATACAC----TAAGGAAACTCAGAAAGC 236

Qy      3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db       237 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCAGGCCCTAAAGAA 296

Qy      3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db       297 GGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCTTTATAT 354

Qy      3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db       355 GTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGACCAGCTT 413

Qy      3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db       414 GCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 473

Qy      3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db       474 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAATACAGGG 533

Qy      3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689

Db       534 AAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 593

Qy      3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db       594 CTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGGGCCA 651


RESULT 433
AB115221
LOCUS       AB115221                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP9-1.
ACCESSION   AB115221
VERSION     AB115221.1  GI:41016651
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP9-1"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.7;  DB 1;  Length 685;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  563;  Conservative   59;  Mismatches   32;  Indels   35;  Gaps    5;

Qy      3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136

Db         1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy      3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196

Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy      3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251

Db        121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy      3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298

Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy      3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358

Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy      3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418

Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy      3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478

Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy      3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538

Db        417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy      3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598

Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy      3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658

Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy      3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718

Db        597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy      3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 434
AB115304
LOCUS       AB115304                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP17-5.
ACCESSION   AB115304
VERSION     AB115304.1  GI:41052532
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
```

```
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP17-5"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.7;  DB 1;  Length 685;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  563;  Conservative   59;  Mismatches   32;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
             |||||||:|:|:|||||:|||||||||||||||                |||||||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
             :||||:||||||||||||||||:||||||||||||||:::||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
             ||||||| |||:||||||||||| |||||||: ::|||||||||:|||:|||:
Db         121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                         |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db         181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db         240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             |||:::| |  ||||||||||||    ||  |   |||  |||||||||:||||||||||
Db         300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :|||||| |:|||||||||||||:||: ||::|||||| |||||| ||:||:|||||:|
Db         358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||||||:|||:|||||: ||||||||||||::||||::|||||||:||||||||
Db         417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
             :|||||||||||||||||||:|:||||||||||:|||||||||| ||||||||||||::
Db         477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
             |||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db         537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
             |:|||||: |||||||||||||||||| ||||| |||||||||||||||:|||| ||||
Db         597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||:||||||||||||||||||:|||||
Db         657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 435
AB126123
```

```
LOCUS       AB126123                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP7-2.
ACCESSION   AB126123
VERSION     AB126123.1  GI:47169723
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP7-2"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.6;  DB 1;  Length 638;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   33;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||  |||||||:|:: |||||||||:|||:|||:|||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||: |||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:|||||||||:|:||||||:::||| |||||||||||||  ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||| :|||||  |:|||||||||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAAGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             : ||||| ||||||||||:||:||||||:||||||||||||||:|| :|||||:||||||
Db         342 GCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:||||||||:|||||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||||  ||::|||||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520
```

```
Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:||||::|||| || |||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:|||||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAAGGCCA 638


RESULT 436
AB126126
LOCUS       AB126126                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP8-1.
ACCESSION   AB126126
VERSION     AB126126.1  GI:47169726
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP8-1"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.6;  DB 1;  Length 638;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   33;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||  |||||||:|:: |||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:|||||:|||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||   ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||  |:|||||  |||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:|||||||||||||||||||:|| ||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGTCA 638


RESULT 437
AB126128
LOCUS       AB126128                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP8-7.
ACCESSION   AB126128
VERSION     AB126128.1  GI:47169728
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP8-7"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.6;  DB 1;  Length 638;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   33;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||  |||||||:|:: |||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:|||||:|||||:|||||||
```

```
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||:|||||||||:|:||||||:::||| ||||||||||||||   ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||| :|||||| |:|||||||||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAAGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              : ||||| |||||||||:||:|||||||:||||||||||||||:|| :|||||:||||||
Db        342 GCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:||||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 438
AB126129
LOCUS       AB126129                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP9-1.
ACCESSION   AB126129
VERSION     AB126129.1  GI:47169729
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP9-1"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.6;  DB 1;  Length 638;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   33;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| ||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||| |||||||:|:: |||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:||||||:||||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||:|||||||||:|:||||||:::||| ||||||||||||||   ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||| :|||||| |:|||||||||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAAGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              : ||||| |||||||||:||:|||||||:||||||||||||||:|| :|||||:||||||
Db        342 GCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:||||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 439
AB126131
LOCUS       AB126131                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP9-3.
ACCESSION   AB126131
VERSION     AB126131.1  GI:47169731
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP9-3"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
```

```
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.6;  DB 1;  Length 638;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   33;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| ||||||||||||||||||||:||||:|||||||
Db             1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db            61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||  |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db           121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||| ||: |||: |||||||||||:|||||:|||||:|||||||
Db           181 -CCCAAGTACGGCAGAATAGCCAGGCCATTATATATACACTAATTAA--------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:||||||||:|:||||||:::|||||||||||||||||  ||  |
Db           225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||| |:|||||| |||||||:||:|||:
Db           284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||:|| :|||||:||||||
Db           342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db           401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||||  ||::|||||||||||||||||||||||| | ||
Db           461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||||||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db           521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||: || |||||||:|||||||||||||||||:|||||
Db           581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 440
BV486829/c
LOCUS       BV486829                671 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S215P60148FD5.T0 Clara Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV486829
VERSION     BV486829.1  GI:62330499
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 671)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 671
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .671
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Clara"
     STS             <1. .>671

  Query Match             6.8%;  Score 511.5;  DB 1;  Length 671;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  563;  Conservative   62;  Mismatches   42;  Indels    7;  Gaps    6;

Qy       2865 AACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCAT 2924
              ||| ||:  |||||  ||||||||||:||||| |||| |  ||||||||||:|| |||
Db            668 AACTCAATGTCTCACTTCACCTGGACTGTTTTA-CCCAGGATTTCAGGGATAGCCCACAT 610

Qy       2925 CTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTC 2984
              |||||||||||||||||||| |||||||    : | |::|||||||||||||||||||||
Db            609 CTATTTGGCCAGGCATTAGCTCAAGACTGAGCCAATTTCTCATACCTGGACACTCTTGTC 550

Qy       2985 CTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCC 3044
              ||||:|||:|||||||||||||| ||:||: ||::|||||||||||||||||||||||||
Db            549 CTTCAGTAGGTGGATGATTTACTGTTAGCCACCCGTTCAGAAACCTTGTGCCATCAAGCC 490

Qy       3045 ACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCT 3104
              ||||||| :|||||:||||||||||:||||||||||||:||||||||||:|:|:|||
Db            489 ACCCAAGTGCTCTTAAATTTCCTCGCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCT 430

Qy       3105 CARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCC 3164
```

```
                ||:|||| |||||||||||||||||||||||:||| : |||||| |||||||||:||||
Db        429 CAGCTCTTCTCACAGCAGGTTAAATACTTAGGGCTACAGTTATCCCAAGGCACCAGGGCC 370

Qy       3165 CTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAA 3224
                |||||||||||:::|||||||||||||||||||||||||||:|||||||||||||||
Db        369 CTCAGTGAGGAATGTATCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAAGCAA 310

Qy       3225 CTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRA 3284
                |||||:|::|||||||||:|||:|||:||||||| ||:|||||||| |||| ||:  ||:|
Db        309 CTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCAAATATGGATT-CCCAAGTACAGCAA 251

Qy       3285 AATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTART 3344
                ||||||||| :|||||:|||||:|||||||||||||||  |||  | |   |||||||:|
Db        250 AATAGCCAGACCATTATATACACTAATTAAGGAAACTC--AAAAGCCAATACCATTTAGT 193

Qy       3345 AAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAA 3404
                ||||||||:|:||||||:::|||  |||||||||||  |||||||||||||| |||||||
Db        192 AAGATGGACACCTGAAGCAGAAGCAGCTTTCCAGGC-ACTAAAGAAGGCCTT-AACCCAA 135

Qy       3405 GYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAC 3464
                |:||||||||||||:||||||||:|||||||||||||:||:|||::|||||  ||||||
Db        134 GCCCCAGTGTTAAGTTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-GAAAAAC 76

Qy       3465 AGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYA 3524
                ||:|| ||||||:|||||| ||||||||:|||:| |||:||||||||||::|| |::|
Db         75 AGGAAGAGCTCTAGGAGTCTTTACACAGGTCCGAAGGACGAGCTTGCAACCCATGCCACA 16

Qy       3525 CCTGASTAAGGAAA 3538
                |||||:||||||||
Db         15 CCTGAGTAAGGAAA 2


RESULT 441
AB114928
LOCUS       AB114928                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-17.
ACCESSION   AB114928
VERSION     AB114928.1  GI:41016436
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-17"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 511.2;  DB 1;  Length 540;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  519;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                | | ||||| || |||||||||| ||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                |||||||||||||   |||||||||||||||||| ||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 442
AB114967
LOCUS       AB114967                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-11.
ACCESSION   AB114967
VERSION     AB114967.1  GI:41016475
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-11"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 511.2;  DB 1;  Length 540;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  519;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||   ||||||||||||||||||| ||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 443
AB114923
LOCUS       AB114923                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-12.
ACCESSION   AB114923
VERSION     AB114923.1  GI:41016431
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-12"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .537
                     /note="HERV-W env"
```

```
  Query Match             6.8%;  Score 511.1;  DB 1;  Length 537;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches  525;  Conservative    4;  Mismatches    8;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||| ||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGGAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 444
AB115248
LOCUS       AB115248                658 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP1-2.
ACCESSION   AB115248
VERSION     AB115248.1  GI:41052476
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 658)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP1-2"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
```

```
                     /note="HERV-W pol"
     repeat_region   1. .658
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 511.1;  DB 1;  Length 658;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  555;  Conservative   58;  Mismatches   42;  Indels    7;  Gaps    5;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTAT- 3147
Db          1 TCCAAACCAAAAAGTTCAGCTTTGCTTTCAGCAGGTTTAATACTTAGGGGCTAAAATTAT 60

Qy       3148 -CCAAAGGCACCARGGCCCTCAGTGAGG-AAYRYATCCAGCCTATACTGGCTTATCCTCA 3205
Db         61 CCCAAAGGCACCAGGGCCCTCAGTGAGGAAACATATCCAGCCTATACTGGCTTATCCTCA 120

Qy       3206 TCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATG 3265
Db         121 TCCCAAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATG 180

Qy       3266 GATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGA 3325
Db         181 GATT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGA 239

Qy       3326 AAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
Db         240 AAGTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAA 299

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
Db         300 AGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTT 357

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
Db         358 ATATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGA 416

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
Db         417 GCTTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTC 476

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
Db         477 ATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATAC 536

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
Db         537 AGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAG 596

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
Db         597 GAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGC 656

Qy       3746 CA 3747
Db         657 CA 658


RESULT 445
AB115285
LOCUS       AB115285                 658 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP9-3.
ACCESSION   AB115285
VERSION     AB115285.1  GI:41052513
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 658)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP9-3"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W pol"
     repeat_region   1. .658
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.7;  DB 1;  Length 658;
  Best Local Similarity   84.7%;  Pred. No. 0;
  Matches  561;  Conservative   58;  Mismatches   36;  Indels    7;  Gaps    6;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGG-RCTAARATTAT 3147
Db          1 TCCAACCCAAAAAGTTCAGCTCTGCTTTCAGCAGGTTTAATACTTAGGGGCTAAAATTAT 60

Qy       3148 CCAAA-GGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
Db         61 CCAAAGGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCAT 120

Qy       3207 CYCAAAACCCTAAAGCAACTAAG-RGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATG 3265
Db         121 CCCAAAACGCTAAAGCAACTAAGAGGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATG 180

Qy       3266 GATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGA 3325
Db         181 GATT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGA 239

Qy       3326 AAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
Db         240 AAGTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAA 299

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
Db         300 AGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTT 357

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
Db         358 ATATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGA 416

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
Db         417 GCTTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTC 476

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
Db         477 ATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATAC 536

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
Db         537 AGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAAG 596

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
Db         597 GAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGC 656

Qy       3746 CA 3747
Db         657 CA 658
```

```
RESULT 446
AB115261
LOCUS       AB115261                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP4-1.
ACCESSION   AB115261
VERSION     AB115261.1  GI:41052489
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP4-1"
                     /db_xref="taxon:9606"
                     /cell_line="HCT-116"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.7;  DB 1;  Length 685;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   32;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
              |||||||:|:|:|||||:|||||||||||||||              ||||||||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
              :||||:||||||||||||||||:||||||||||||||:::|||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
              ||||||| |||:|||||||||||| ||||||||: ::|||||||||:|||:|||:
Db         121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                         |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db         181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
              ||:||||:|||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db         240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
              |||:::| |  |||||||||||    ||  |   |||  ||||||||:|||||||||||
Db         300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
              :|||||  |:|||||||||||||:||: ||::|||||| |||||| ||:||:|||||:|
Db         358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
              |||||||||||||:|||:|||||: ||||||||||::|||||::|||||:||||||||
Db         417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
              :||||||||||||||||||||:|:||||||||||:||||||||| ||||||||||||::
Db         477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
              |:|||||: |||||||||||||||||| ||||| |||||||||||||||:|||| ||||
Db         597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCC 3746
              |||:|||||||||||||||||||:||||
Db         657 AATATCAGGCTCTATTACTTGAAGGGCC 684


RESULT 447
AB115199
LOCUS       AB115199                668 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP3-4.
ACCESSION   AB115199
VERSION     AB115199.1  GI:41016629
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 668)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP3-4"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .668
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.6;  DB 1;  Length 668;
  Best Local Similarity   82.9%;  Pred. No. 0;
  Matches  557;  Conservative   59;  Mismatches   38;  Indels   18;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
              |||||||:|:|:|||||:|||||||||||||||              ||||||||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
              :||||:||||||||||||||||:||||||||||||||:::|||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTG 3256
              ||||||||||:|||||| ||||||||||||||:|::||||||||:|||:|||:|||||
Db         121 TTATCCTCATCCCAAAACTCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTG 180

Qy       3257 CCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGG 3316
               ||||:||||||| |||||||: || :|||||| ||| :|||| :|||||:||||||||
Db         181 TCGAATATGGATT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGG 239

Qy       3317 AAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCC 3376
```

```
                |||||  |||||  |||||||||||||:|||||||||:|:||||||:::||| ||||||||
Db        240 AAACTTGGAAAGTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCC 299

Qy       3377 AGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGA 3436
                ||||||  ||  | | |||  ||||||||:|||||||||||||:|||||  | :||||||||
Db        300 AGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGA 357

Qy       3437 CTTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCC 3496
                |||||:||:|||::||| | ||||||| |||:||:|||||||:|||||||||||||||:|||
Db        358 CTTTTCTTTATATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCC 416

Qy       3497 RAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGG 3556
                :|||||:||||||||||||::|||||::|||||| ||||||||:|||||||||||||||||
Db        417 GAGGGACGAGCTTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGG 476

Qy       3557 GTTGRCYTCATTGTTTAY-GGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT 3615
                ||||:|:||||||||||: |||||||||  ||||||||||::|||||||||||||||||||
Db        477 GTTGGCCTCATTGTTTATGGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTT 536

Qy       3616 AAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTC 3675
                |||||||||||||:|||||| |||||| ||||||||||||||:| |||::| |||||
Db        537 TAAATAATACAGGGAAGAGATTTTACTGGGTGGACATCTCATGATGGGAACGGGATACTG 596

Qy       3676 ACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTA 3735
                |||||||||||| ||||| ||||| || |||||:| ||||||||||:| ||||||||||
Db        597 ACTGCTAAAGGAAACTTGGGGCTGGCAAACAACCGGTTACTTAAATATAAGGCTCTATTA 656

Qy       3736 CTTGAARGGCCA 3747
                ||||||:|||||
Db        657 CTTGAAGGGCCA 668


RESULT 448
BD003153
LOCUS       BD003153                 542 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Secretory proteins and polynucleotides encoding the same.
ACCESSION   BD003153
VERSION     BD003153.1  GI:18631114
KEYWORDS    JP 2001501813-A/33.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Mccoy,J.M., Levallie,E.R., Racie,L.A., Merberg,D., Treacy,M.,
            Evans,C. and Spaulding,V.
  TITLE     Secretory proteins and polynucleotides encoding the same
  JOURNAL   Patent: JP 2001501813-A 33 13-FEB-2001;
            GENETICS INSTITUTE INC,KENNETH JACOBS
COMMENT     OS   Unidentified
            PN   JP 2001501813-A/33
            PD   13-FEB-2001
            PF   23-JUL-1997 JP 1998508939
            PR   26-JUL-1996 US    08/686878
            PI   JOHN M MCCOY,EDWARD R LEVALLIE,LISA A RACIE,DAVID MERBERG, PI
             MAURICE TREACY,
            PI   CHERYL EVANS,VIKKI SPAULDING
            PC   C12N15/09,A61K38/00,A61K48/00,A61P3/02,A61P7/02,A61P7/08, PC
            A61P35/00,
            PC   A61P37/00,C07K14/47,C12N5/10,C12P21/02,C12N15/00,C12N5/00, PC
             A61K37/02
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .542
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             6.8%;  Score 510.3;  DB 1;  Length 542;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  516;  Conservative   12;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
                ||||| || |||||||||||||||||||| ||||| |||| |||||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
                ||||::|||||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
Db         61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy       5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
                | ||||| ||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db        121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy       5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy       5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy       5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
                |||||||||||||||||||||||||||||||||||| |||||||:|||||||||||||
Db        301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy       5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
                ||||||||||:||:||||:|||:|||||:|||||||||||||||||||||||||||||||
Db        360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy       5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
                ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
Db        420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy       5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
                ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy       5779 CAT 5781
                |||
Db        540 CAT 542


RESULT 449
I80072
LOCUS       I80072                   542 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 48 from patent US 5708157.
ACCESSION   I80072
VERSION     I80072.1  GI:3208362
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Evans,C. and Spaulding,V.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5708157-A 48 13-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             6.8%;  Score 510.3;  DB 1;  Length 542;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  516;  Conservative   12;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5239 AACAAGTTCTTAAAACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCC 5298
                ||||| || |||||||||||||||||||| ||||| |||| |||||||||||||| ||||||
Db          1 AACAANTTNTTAAAACATTACAAGGAACGTATCCNTGAGNAGAGGGAAAAGAATTATTCC 60

Qy       5299 ACCCWWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACAT 5358
                ||||::|||||||||||||||||||||| |||:|| |||||||| ||||||||| ||||
```

```
Db        61 ACCCTTGTGACATGGTATTAGTCAAGTCCATTCCCTATAATTCCCNATCCCTAGANACAT 120

Qy      5359 CCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAG 5418
             | ||||| ||||| ||||||||||||||| |||||||||||||||||||||||||||||
Db       121 CNTGGGANGGACCNTACCCAGTCATTTTATNTACCCCAACTGCGGTTAAAGTGGCTGGAG 180

Qy      5419 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 5478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAA 240

Qy      5479 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 5538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCA 300

Qy      5539 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCCTTATCAT 5598
             |||||||||||||||||||||||||||||||||| ||||||||:|||||||||||||||
Db       301 AACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCCTTATCAT 359

Qy      5599 ATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCCATGCCGC 5658
             ||||||||||:||:||||:|||:|||||:||||||||||||||||||||||||||||||
Db       360 ATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCCATGCCGC 419

Qy      5659 TGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 5718
             ||||||||||||||||||||:|||:||||||||||||||||||||||||||||||||||
Db       420 TGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAAT 479

Qy      5719 ATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGCCCACACC 5778
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db       480 ATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACC 539

Qy      5779 CAT 5781
             |||
Db       540 CAT 542
```

```
RESULT 450
AB115188
LOCUS       AB115188                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP1-4.
ACCESSION   AB115188
VERSION     AB115188.1  GI:41016618
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP1-4"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.1;  DB 1;  Length 685;
  Best Local Similarity   81.6%;  Pred. No. 0;
```

```
  Matches  562;  Conservative   59;  Mismatches   33;  Indels   35;  Gaps    5;

Qy      3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG----------GTTAAATACTTAGG 3136
             |||||||:|:|:|||||:|||||||||||||||              ||||||||||||
Db         1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy      3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
             :||||:||||||||||||||||:||||||||||||||:::|||||||||||||||||||
Db        61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy      3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
             ||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db        121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy      3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                         |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy      3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
             ||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy      3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
             |||:::| |  |||||||||||    ||  |   |||  ||||||||:||||||||||||
Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy      3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
             :|||||| |:|||||||||||||:||: ||::|||||| |||||| ||:||:||||||:|
Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy      3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
             |||||||||||||:|||:||||||||||||||::||||::||||||:||||||||||||
Db        417 GAGTCCTTACACAGGTCCGAGGGACGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy      3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
             :|||||||||||||||||||:|:||||||||||:||||||||| ||||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy      3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
             |||||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy      3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
             |:|||||: |||||||||||||||||| ||||| ||||||| |||||||||:| || ||||
Db        597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGGCAGACAACTGGTTGCTTA 656

Qy      3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||:||||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685
```

```
RESULT 451
AB115192
LOCUS       AB115192                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP2-4.
ACCESSION   AB115192
VERSION     AB115192.1  GI:41016622
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
```

```
          Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
          of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
          Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP2-4"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.1;  DB 1;  Length 685;
  Best Local Similarity   81.6%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   33;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
                ||||||:|:|:|||||:|||||||||||||||            ||||||||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
                :||||:||||||||||||||||:||||||||||||||||:::||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
                ||||||| |||:|||||||||||| ||||||||: ::||||||||||:|||:|||:
Db        121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                         |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
                ||:|||||:|||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
                |||:::| |  |||||||||||  ||  |   |||  ||||||||:||||||||||||
Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
                :|||||| |:|||||||||||:||: ||::|||||| |||||| ||:||:|||||:|
Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
                |||||||||||||:|||:|||||: ||||||||||::||||::||||||:||||||||
Db        417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
                :|||||||||||||||||||||:|:||||||||||:|||||||| | ||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGACAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
                |||||||||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
                |:|||||: |||||||||||||||||| ||||| |||||||||||||||:|||| ||||
Db        597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
                |||:|||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 452
AB115237
LOCUS       AB115237                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP11-8.
ACCESSION   AB115237
VERSION     AB115237.1  GI:41016667
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP11-8"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510.1;  DB 1;  Length 685;
  Best Local Similarity   81.6%;  Pred. No. 0;
  Matches  562;  Conservative   59;  Mismatches   33;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
                ||||||:|:|:|||||:|||||||||||||||            ||||||||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
                :||||:||||||||||||||||:||||||||||||||||:::||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251
                ||||||| |||:|||||||||||| ||||||||: ::||||||||||:|||:|||:
Db        121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3252 ---------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
                      |||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
                ||:|||||:|||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
                |||:::| |  |||||||||||  ||  |   |||  ||||||||:||||||||||||
Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
                :|||||| |:|||||||||||:||: | ::|||||| |||||| ||:||:|||||:|
Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
                |||||||||||||:|||:|||||: ||||||||||::||||::||||||:||||||||
Db        417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
                :|||||||||||||||||||||:|:||||||||||:|||||||| | ||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
                |||||||||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596
```

```
Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
              |:||||: |||||||||||||||||| ||||| |||||||||||||||:|||| ||||
Db        597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||:||||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 453
AB126101
LOCUS       AB126101                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP1-2.
ACCESSION   AB126101
VERSION     AB126101.1  GI:47169701
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP1-2"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510;  DB 1;  Length 638;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  545;  Conservative   59;  Mismatches   34;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||||:||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||  |||||||:|:: |||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: |||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||| |:|||||  |||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341
```

```
Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||||:|| :|||||:||||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||||||||| |||||||:|||||::|||| || |||||||||| ||||
Db         521 AAGAGATGTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||| ||: || ||||||||:|||||||||||||||||:|||||
Db         581 CTTGTGGCTGTCAGACCACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 454
AB126105
LOCUS       AB126105                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP2-2.
ACCESSION   AB126105
VERSION     AB126105.1  GI:47169705
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP2-2"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 510;  DB 1;  Length 638;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  545;  Conservative   59;  Mismatches   34;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              |||||||||:||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||  |||||||:|:: |||||||||:|||:|||:||||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
```

```
                 ||||||| ||| |||||||||| ||||| ||||| |||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                 |||||||||||| |||||||||| ||||||:::|||||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                 | |||  |||||||| |||||||||||  |||||| |||||  |||||||:|||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
                 :|||||| |||||||||:||:|||||||:|||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
                 ||||||::|||| :|||| |:||||||||:|||||||||||||| ||||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
                 ||||: |||||||| ||||||||| ||::|||||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
                 :|||||| |||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATGTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                 |||||||||||||||| ||: || |||||||:|||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACCACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638
```

RESULT 455
AB126106
LOCUS AB126106 638 bp RNA linear PRI 14-MAY-2004
DEFINITION Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP2-11.
ACCESSION AB126106
VERSION AB126106.1 GI:47169706
KEYWORDS .
SOURCE Macaca fuscata (Japanese macaque)
ORGANISM Macaca fuscata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS Yi,J., Takenaka,O. and Kim,H.
TITLE Expression and identification of HERV-W family in Japanese Monkeys (Macaca fuscata)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 638)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .638
/organism="Macaca fuscata"
/mol_type="other RNA"
/isolate="JM-WP2-11"
/db_xref="taxon:9542"
/tissue_type="prostate"
/note="HERV-W pol"
repeat_region 1. .638
/note="HERV-W pol"

Query Match 6.8%; Score 510; DB 1; Length 638;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 545; Conservative 59; Mismatches 34; Indels 20; Gaps 4;

```
Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
                 |||||||:|:|:|||||:||||| |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
                 ||||||||||:|||||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
                 |||||||||||  |||||||:|:: |||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                 |||||||: |||: |||||||||:|||||:|||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                 |||||||||||| |||||||||| |:|||||:::|||||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                 | |||  |||||||| :|||||||||||  :|||||  |:|||||  |||||||:|||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
                 :|||||| |||||||||:||:|||||||:|||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
                 ||||||::|||| :|||| |:||||||||:|||||||||||||||| ||||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
                 ||||: |||||||| ||||||||| ||::|||||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
                 :|||||| |||||||| ||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATGTTACTGTGCGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
                 |||||||||||||||||: || |||||||:|||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638
```

RESULT 456
AB126109
LOCUS AB126109 638 bp RNA linear PRI 14-MAY-2004
DEFINITION Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-3.
ACCESSION AB126109
VERSION AB126109.1 GI:47169709
KEYWORDS .
SOURCE Macaca fuscata (Japanese macaque)
ORGANISM Macaca fuscata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS Yi,J., Takenaka,O. and Kim,H.
TITLE Expression and identification of HERV-W family in Japanese Monkeys (Macaca fuscata)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 638)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .638
/organism="Macaca fuscata"
/mol_type="other RNA"
/isolate="JM-WP5-3"
/db_xref="taxon:9542"

```
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match              6.8%;  Score 510;  DB 1;  Length 638;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  545;  Conservative   59;  Mismatches   34;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||  |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:|||||:|||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||| |:|||||| |||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||:|| :||||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||| ||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATGTTACTGTGCGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||: || ||||||||:||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 457
AB126111
LOCUS       AB126111                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-6.
ACCESSION   AB126111
VERSION     AB126111.1  GI:47169711
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-6"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match              6.8%;  Score 510;  DB 1;  Length 638;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  545;  Conservative   59;  Mismatches   34;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||  |||||||:|:: ||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:|||||:|||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||| |:|||||| |||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||:|| :||||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::|||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||||||||| ||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATGTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||: || ||||||||:||||||||||||||||:|||||
Db        581 CTTGTGGCTGTCAGACCACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 458
AB115097
LOCUS       AB115097                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE1-4.
ACCESSION   AB115097
VERSION     AB115097.1  GI:41016532
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE1-4"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 509.6;  DB 1;  Length 540;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  518;  Conservative    4;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||| || |||||||||||||||||||||||| |||||||||||||||||| |||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACTCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 459
AB115238
LOCUS       AB115238                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP11-9.
ACCESSION   AB115238
VERSION     AB115238.1  GI:41016668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP11-9"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 509.1;  DB 1;  Length 685;
  Best Local Similarity   81.5%;  Pred. No. 0;
  Matches  561;  Conservative   59;  Mismatches   33;  Indels   35;  Gaps    5;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGGR 3137
              ||||||:|:|:||||||:|| ||||||||||||               ||||||||||||:
Db          2 CCAAACCAAAGGCTCAGCTTTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGGG 61

Qy       3138 CTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCT 3197
              ||||:||||||||||||||:||||||||||||||:::|||||||||||||||||||||||
Db         62 CTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCT 121

Qy       3198 TATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT------------------ 3239
              |||||| |||:|||||||||||| ||||||||: ::||||||
Db        122 TATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAACA 181

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
              |||:|||:||||:|||||| |||:|||||||||| | ||: |||:|||||||||| :||||
Db        182 GGCATAACAGGTTTCTGCTGAATATGGATTCCCAA-GTACGGCAAAATAGCCAGACCATT 240

Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGA 3359
              |:|||||:|||||||||||||||||||||||||||||||||:|||||||||:|:||||
Db        241 ATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGA 300

Qy       3360 AGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGY 3419
              ||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||:
Db        301 AGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAGC 358

Qy       3420 TTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGG 3479
              |||||| |:|||||||||||:||: ||::|||||| |||||| ||:||:|||||:||
Db        359 TTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAGG 417

Qy       3480 AGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAY 3539
              |||||||||||:|||:|||||: ||||||||||::||||::||||||:||||||||:
Db        418 AGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAAT 477

Qy       3540 TGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKT 3599
              |||||||||||||||||||:|:||||||||:|||||||| ||||||||||||::|
Db        478 TGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCTT 537

Qy       3600 AGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGA 3659
```

```
                |||||||||||||||||||||||||||| ||:|||||||||||||||||| || ||||||||
Db        538 AGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATGA 597

Qy       3660 KGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAA 3719
              :|||||: |||||||||||||||||| ||||| ||||||||||||||:|||| |||||
Db        598 TGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTAA 657

Qy       3720 ATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||:|||||||||||||||||:|||||
Db        658 ATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 460
AB115113
LOCUS       AB115113                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE5-2.
ACCESSION   AB115113
VERSION     AB115113.1  GI:41016548
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE5-2"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 508.4;  DB 1;  Length 540;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  518;  Conservative    3;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        421 CACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 461
AB126110
LOCUS       AB126110                 638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP5-5.
ACCESSION   AB126110
VERSION     AB126110.1  GI:47169710
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP5-5"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 508.4;  DB 1;  Length 638;
  Best Local Similarity   82.7%;  Pred. No. 0;
  Matches  544;  Conservative   59;  Mismatches   35;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||:::|||||||| ||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||  |||||||:|:: ||||||||:|||:|||:|||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: |||||||||||:||||||:|||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||  ||  |
```

```
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||| :|||||| |:||||| ||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:|||||||:||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||||| |||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::||||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||| ||||| |||||||:|||||::|||| || |||||||| ||||
Db        521 AAGAGATGTTACTGTGCGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              | |||||||||||||||||: || ||||||||:|||||||||||||||||:|||||
Db        581 CCTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 462
AB126121
LOCUS       AB126121                638 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP6-9.
ACCESSION   AB126121
VERSION     AB126121.1  GI:47169721
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 638)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP6-9"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W pol"
     repeat_region   1. .638
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 508.4;  DB 1;  Length 638;
  Best Local Similarity   82.7%;  Pred. No. 0;
  Matches  544;  Conservative   59;  Mismatches   35;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||  |||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||  |||||||:|:: ||||||||:|||:|||:|||||||| |:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGGATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: ||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:||||||||:|:||||||:::|||||||||||||||||||   ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||| :|||||| |:||||| ||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:|||||||:||||||||||||||:|| :|||||:||||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:||||||||:|||||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| |||||||| ||::||||||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||| ||||| |||||||:|||||::|||| || |||||||| ||||
Db         521 AAGAGATGTTACTGTGCGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||||||||||||||||: || ||||||||:|||||||||||||||||:|||||
Db         581 CTTGTGGCTGTCAGACAACCATTCACTTAAATATCAGGCTCTATTACTTGAAGGGCCA 638


RESULT 463
AB056482
LOCUS       AB056482                642 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP2-3.
ACCESSION   AB056482
VERSION     AB056482.1  GI:15341210
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 642)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP2-3"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .642
                     /gene="pol"
```

```
     misc_feature    <1. .>642
                     /gene="pol"

  Query Match             6.8%;  Score 508.4;  DB 1;  Length 642;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  552;  Conservative   57;  Mismatches   33;  Indels   16;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:||||| |||||||||||||||||||||||: |||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGTTCACAGCAGGTTAAATACTTAGGGTTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||| ||||||||:::||||| ||| |||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCTGTGAGGAATGTATCCAACCTGTACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||: :: |||||||||:|||:|||:|||||| |||: ||||||
Db        121 AAAACCCTAAAGCAACTAAGAAGG-TCCTTGGCATAACAGGTTTCTGCTGAATGTGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
               ||||||:   |:|||||||||| :||||:|||||:|||||| ||||||||||||||
Db        180 -TCCAGGTACAACGAAATAGCCAGACCATTATATACACTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:||||||:::|||  |||||||||| ||  || ||
Db        239 CAATACCCATTTAGTAAGATGGATACCTGAAGCAGAAGCAGCTTTCCAGGGCCTAAAGGA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
               | |||  ||||||||:|||||||||||:|||||||: |||||||||||:||:|||:
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGCGGCAAGACTTTTCTTTATAT 356

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||| ||:||:|||||:|||||||||||||||:|||:| |||:|||||||
Db        357 GTCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAAGGATGAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||           :||||||:||||||||:||||| |||||||||||||||:|:||||||
Db        416 GC-----------CACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGTTGGCCTCATTG 464

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:|||||||| ||||||||||::|||||||||||||||||||||||||||||||||||
Db        465 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 524

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|| ||| |||||||||||||| ||||||||:|||||::|| ||||||||||||||||||||
Db        525 AAGGGATATTACTGTGTGGACGTCTCATGATGTGAATGGCGTACTCACTGCTAAAGGAGA 584

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||: ||||||||||:|||||||||||| |||||:|||||
Db        585 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTATTTGAAGGGCCA 642


RESULT 464
AB115305
LOCUS       AB115305                686 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP18-1.
ACCESSION   AB115305
VERSION     AB115305.1  GI:41052533
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 686)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP18-1"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W pol"
     repeat_region   1. .686
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 507.9;  DB 1;  Length 686;
  Best Local Similarity   81.3%;  Pred. No. 0;
  Matches  561;  Conservative   59;  Mismatches   35;  Indels   35;  Gaps    5;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAG-------------GTTAAATACTTAG 3135
             | | || |:|:|:|||||:|||||||||||||||              ||||||||||
Db          1 TCCAAACCCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAG 60

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:||||:|||||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db         61 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 120

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT---------------- 3239
             |||||||| |||:|||||||||| ||||||||: ::||||||
Db        121 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 180

Qy       3240 --GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                |||:|||:|||:|||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db        181 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 239

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        240 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 299

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||    ||  |  |||  |||||||:|||||||||||
Db        300 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 357

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: ||::|||||| |||||| ||:||:||||||:
Db        358 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTA 416

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||:|||:|||||: ||||||||||::||||::||||||:|||||||
Db        417 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 476

Qy       3538 AYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTY 3597
             |:|||||||||||||||||:|:||||||||:|||||||| |||||||||:
Db        477 ATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTC 536

Qy       3598 KTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCAT 3657
             :|||||||||||||||||||||||||||| ||:|||||||||||||||| || ||||||
Db        537 TTAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCAT 596

Qy       3658 GAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTT 3717
             ||:|||||: |||||||||||||||||| ||||| |||||||||||||||:|||| |||
Db        597 GATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTT 656

Qy       3718 AAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||:|||||||||||||||||||:|||||
Db        657 AAATATCAGGCTCTATTACTTGAAGGGCCA 686


RESULT 465
AB114931
LOCUS       AB114931                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-20.
ACCESSION   AB114931
```

```
VERSION     AB114931.1  GI:41016439
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-20"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.8%;  Score 507.7;  DB 1;  Length 539;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches  523;  Conservative    4;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 466
AB114960
```

```
LOCUS       AB114960                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-1.
ACCESSION   AB114960
VERSION     AB114960.1  GI:41016468
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-1"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.8%;  Score 507.7;  DB 1;  Length 539;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches  523;  Conservative    4;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || ||||||||| |||| || ||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAAC-AAGGAGTACCCATTCTTCCTTTTGTTATCAGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||   ||||||||||||||||||| ||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539
```

```
RESULT 467
AB126102
LOCUS       AB126102                639 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP1-4.
ACCESSION   AB126102
VERSION     AB126102.1  GI:47169702
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 639)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP1-4"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .639
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 507.4;  DB 1;  Length 639;
  Best Local Similarity   82.6%;  Pred. No. 0;
  Matches  543;  Conservative   59;  Mismatches   35;  Indels   20;  Gaps    4;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:||||| |||||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              ||||||||||  |||||||:|:: |||||||||:|||:|||:|||||||||||:|||||||
Db         121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||: |||||||||||:|||||:|||||:|||||||
Db         181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:|:||||||:::|||||||||||||||||||   ||  |
Db         225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:|||||||||||| :|||||| |:|||||  |||||||:||:|||:
Db         284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||||:|| :|||||:||||||
Db         342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::|||| :|||| |:|||||||||:|||||||||||||| |||||:|:||||||
Db         401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
              ||||: |||||||| ||||||||| ||::|||||||||||||||||||||||| | ||
Db         461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
              :|||||| |||||||||||||| |||||||:|||||::|||| || |||||||| ||||
Db         521 AAGAGATGTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTACTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
              |||||||||||||||| ||: || ||||||||:|||||||||||||||||:|| |
Db         581 CTTGTGGCTGTCAGACCACCATTCACTTAAATATCAGGCTCTATTACTTGAAAGGGC 637


RESULT 468
AB054089
LOCUS       AB054089                543 bp    DNA     linear   PRI 06-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:UO31-3.
ACCESSION   AB054089
VERSION     AB054089.1  GI:14780053
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
  TITLE     Molecular cloning and phylogenetic analysis of new human endogenous
            retrovirus HERV-W family in cancer cells
  JOURNAL   Curr. Microbiol. (2001) In press
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Lee,J. and Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="UO31-3"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W env"
     repeat_region   1. .543
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.8%;  Score 507.2;  DB 1;  Length 543;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  523;  Conservative    4;  Mismatches   12;  Indels    2;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db         241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```

```
                ||||||
```

```
Db        61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy      3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269

Db       121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy      3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329

Db       181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy      3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389

Db       235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy      3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449

Db       295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy      3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509

Db       353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy      3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569

Db       412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy      3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629

Db       472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy      3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATA-CTCACTGCTAAAGGAG 3688

Db       532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACCTCACTGCTAAAGGAG 591

Qy      3689 ACTTGTGGCTGTCAGACAACY-GTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746

Db       592 ACTTGTGGCTGTCAGACAACTGGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCC 650


RESULT 471
AB115236
LOCUS       AB115236                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP11-7.
ACCESSION   AB115236
VERSION     AB115236.1  GI:41016666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP11-7"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 506.9;  DB 1;  Length 685;
  Best Local Similarity   81.3%;  Pred. No. 0;
  Matches  560;  Conservative   59;  Mismatches   35;  Indels   35;  Gaps    5;

Qy      3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136

Db         1 TCCAACCCAAAGGCTCAGCTCTGCTCACAGCAGAGGGCTATTAATCCCTAAATACTTAGG 60

Qy      3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196

Db        61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy      3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY----- 3251

Db       121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy      3252 ------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298

Db       181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy      3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358

Db       240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy      3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418

Db       300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy      3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478

Db       358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy      3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538

Db       417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy      3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598

Db       477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy      3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658

Db       537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy      3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718

Db       597 ATGTGAACCGCATACCCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656

Qy      3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db       657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 472
AB115104
LOCUS       AB115104                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE3-1.
ACCESSION   AB115104
VERSION     AB115104.1  GI:41016539
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
```

```
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE3-1"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 506.8;  DB 1;  Length 540;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  517;  Conservative    3;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||   ||||||||||||||||||| ||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 473
AB056481
LOCUS       AB056481                642 bp    DNA     linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens pol gene, HERV-W family, isolate:HWP2-2.
ACCESSION   AB056481
VERSION     AB056481.1  GI:15341209
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.S.
  TITLE     Sequence and phylogeny of HERV-W pol fragments
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (17), 1665-1671 (2001)
   PUBMED   11779355
REFERENCE   2  (bases 1 to 642)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515102259,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HWP2-2"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /note="HERV-W pol
                     human endogenous retrovirus family :HERV-W"
     gene            1. .642
                     /gene="pol"
     misc_feature    <1. .>642
                     /gene="pol"

  Query Match             6.8%;  Score 506.8;  DB 1;  Length 642;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches  551;  Conservative   57;  Mismatches   34;  Indels   16;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||||||||||||||||||||||: |||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGTTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:|||||||| ||||||||:::|||||| ||| |||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCTGTGAGGAATGTATCCAACCTGTACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||||: :: ||||||||:|||:|||:|||||| |||: ||||||
Db        121 AAAACCCTAAAGCAACTAAGAAGG-TCCTTGGCATAACAGGTTTCTGCTGAATGTGGATT 179

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              ||||||:   |:|||||||||| :|||||:|||||:|||||| |||||||||||||||
Db        180 -TCCAGGTACAACGAAATAGCCAGACCATTATATACACTAATTAAAGAAACTCAGAAAGC 238

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:| ||||:::||| |||||||||| || ||
Db        239 CAATACCCATTTAGTAAGATGGATACCCGAAGCAGAAGCAGCTTTCCAGGGCCTAAAGGA 298

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||||||:||||||||: |||||||||||||:||:|||:
Db        299 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGCGGCAAGACTTTTCTTTATAT 356

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||| ||:||:|||||||:|||||||||||||||:|||:| |||:||||||
Db        357 GTCACAG-AAAAAATAGGAATAGCTCTAGGAGTCCTTACACAGGTCCAAAGGATGAGCTT 415

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||          :||||||:||||||||:|||||  |||||||||||||||:|:|||||||
Db        416 GC----------CACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGTTGGCCTCATTG 464

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||||||| |||||||||||:|||||||||||||||||||||||||||||||||
Db        465 TTTACGGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGGG 524

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|| ||| |||||||||||||| ||||||||:|||||::|| ||||||||||||||||||
Db        525 AAGGGATATTACTGTGTGGACGTCTCATGATGTGAATGGCGTACTCACTGCTAAAGGAGA 584

Qy       3690 CTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||: |||||||||||:|||| |||||||| |||||:|||||
Db        585 CTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGCCTCTATTATTTGAAGGGCCA 642
```

```
RESULT 474
AB115230
LOCUS       AB115230                685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP10-4.
ACCESSION   AB115230
VERSION     AB115230.1  GI:41016660
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP10-4"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 506.5;  DB 1;  Length 685;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  559;  Conservative   60;  Mismatches   35;  Indels   35;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
              |||||||:|:|:| |||: | ||||| ||||||              ||||| |||||||
Db          1 TCCAAACCAAAGGGTCAGGTTTGCTCCCAGCAGAGGGCTATTTATCCTTAAAAACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
              :||||:||||||||||||||||:||||||||||||||:::|||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT----------------- 3239
              ||||||| |||:|||| || |||| ||||||||: ::||||||
Db        121 TTATCCTTATCCCAAACCCTTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3240 -GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
              |||:|||:|||:|||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
              ||:|||||:|||||||||||||||||||||||||||||||||:||||||||:|:|||
Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
              |||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||||
Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
              :|||||| |:|||||||||||||:||: ||::|||||| |||||| ||:||:|||||:|
Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
              ||||||||||||||:|||:|||||: |||||||||||::||||::||||||:||||||||
Db        417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
              :|||||||||||||||||||:|:|||||||||:|||||||| ||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
              |||||||||||||||| ||||||||||||:|||||| ||||||||||||||||||||||
Db        537 TAGTATCTGAAGCAGTTTAAATAATACAGGGAAGAGATTTTACTGTGTGGACATCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
              |:|||||::|||||||| ||||||||||||||||||||||||||||||:|||||||||
Db        597 ATGTGAACGGCATACTGACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTA 656

Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||:|||||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 475
AB114969
LOCUS       AB114969                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-15.
ACCESSION   AB114969
VERSION     AB114969.1  GI:41016477
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-15"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 506.4;  DB 1;  Length 540;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches  516;  Conservative    4;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||| |||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
```

```
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 476
AB115136
LOCUS       AB115136                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE10-2.
ACCESSION   AB115136
VERSION     AB115136.1  GI:41016571
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE10-2"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.8%;  Score 506.4;  DB 1;  Length 540;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches  516;  Conservative    4;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAGCCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        241 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 477
AB115211
LOCUS       AB115211                 654 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-3.
ACCESSION   AB115211
VERSION     AB115211.1  GI:41016641
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-3"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
     repeat_region   1. .654
                     /note="HERV-W pol"

  Query Match             6.8%;  Score 506.4;  DB 1;  Length 654;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  556;  Conservative   60;  Mismatches   36;  Indels    8;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149

Db          1 TCCAGACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCA-RGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208

Db         61 AGACACACCAGGGGCCCTCAGTGAGGAATGCATCCAGCCTATACTGGCTTATCCTTATCC 120

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
```



```
Db        121 CAAAACCATAAAGCAATTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGAT 180

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328

Db        181 T-CCCAGGTACAGTGAAATAGCCAGGCCATTATACACACTAATTAAGGAAACTCAGAAAG 239

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388

Db        240 TGAATACCCATTTAGTAAGAAGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGA 299

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448

Db        300 AGGCCCT--AATCCAAGCCCTAGTGTTAAGCTTGCCAAAGGGGCAAGACTTTTCTTTATG 357

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508

Db        358 TGTCACAG-AAAAAACAGGAATAGCTCTAGGATTCCTTACACAGGTCCGAGGGACCAGGT 416

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568

Db        417 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGGCCTCATT 476

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628

Db        477 GTTTATGGGTAATGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAAG 536

Qy       3629 GRAGAGATCTTA-CTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687

Db        537 GAAGAGATCTTATGTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGA 596

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747

Db        597 GACTTGTGGCTGTCAGACAACCATTAGCTTAAATATCAGGCTC--TTATTTGAAGGGCCA 654
```

RESULT 478
AB120045
LOCUS AB120045 520 bp RNA linear VRL 01-FEB-2006
DEFINITION Human endogenous retrovirus W genomic RNA, region corresponding to gag, isolate:HWG1-2.
ACCESSION AB120045
VERSION AB120045.1 GI:45330716
KEYWORDS .
SOURCE Human endogenous retrovirus W
ORGANISM Human endogenous retrovirus W
Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 520)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University, Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .520
/organism="Human endogenous retrovirus W"
/mol_type="genomic RNA"
/isolate="HWG1-2"
/isolation_source="brain of human"
/db_xref="taxon:87786"
misc_feature 1. .520
/note="corresponds to gag"

Query Match 6.8%; Score 505.9; DB 1; Length 520;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 511; Conservative 9; Mismatches 0; Indels 1; Gaps 1;



```
Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691

Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751

Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811

Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871

Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931

Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991

Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051

Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111

Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152

Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520
```

RESULT 479
AB120051
LOCUS AB120051 520 bp RNA linear VRL 01-FEB-2006
DEFINITION Human endogenous retrovirus W genomic RNA, region corresponding to gag, isolate:HWG3-4.
ACCESSION AB120051
VERSION AB120051.1 GI:45330722
KEYWORDS .
SOURCE Human endogenous retrovirus W
ORGANISM Human endogenous retrovirus W
Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 520)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University, Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .520
/organism="Human endogenous retrovirus W"
/mol_type="genomic RNA"
/isolate="HWG3-4"
/isolation_source="testis of human"
/db_xref="taxon:87786"
misc_feature 1. .520
/note="corresponds to gag"

Query Match 6.8%; Score 505.9; DB 1; Length 520;

```
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  511;  Conservative    9;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||:|||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 480
AB120054
LOCUS       AB120054                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG8-3.
ACCESSION   AB120054
VERSION     AB120054.1  GI:45330725
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG8-3"
                     /isolation_source="placenta of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.8%;  Score 505.9;  DB 1;  Length 520;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  511;  Conservative    9;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||:|||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 481
AB120056
LOCUS       AB120056                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG8-6.
ACCESSION   AB120056
VERSION     AB120056.1  GI:45330727
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG8-6"
                     /isolation_source="placenta of human"
                     /db_xref="taxon:87786"
```

```
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.8%;  Score 505.9;  DB 1;  Length 520;
  Best Local Similarity   98.1%;  Pred. No. 0;
  Matches  511;  Conservative    9;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||:||||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTCCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 482
BD196263
LOCUS       BD196263                678 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders.
ACCESSION   BD196263
VERSION     BD196263.1  GI:33006033
KEYWORDS    JP 2002512530-A/29.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
  TITLE     Endogenetic retroviral sequences, associated with autoimmune
            diseases and/or with pregnancy disorders
  JOURNAL   Patent: JP 2002512530-A 29 23-APR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002512530-A/29
            PD   23-APR-2002
            PF   06-JUL-1998 JP 1999508244
            PR   07-JUL-1997 FR     97/08815
            PI   FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
            MANDRAND,
            PI   FRANCOIS MALLET
            PC   C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Endogenetic retroviral sequences, associated with autoimmune
            CC           diseases
            CC   and/or with pregnancy disorders
            FH   Key             Location/Qualifiers
            FT   source          1. .678
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
              ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
              ||||||| | |||||||||| :|||:| |||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
              |||||||||||||||||||||||| :|||||:||||||||||:||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG-----------GTTAAATACTTAG 3135
              |||||||||:|:|:|||||:|||||||||||||||           ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
              |:|| |:|||||||||||||||:|||||||||||||||:::|||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
              |||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
              |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
              ||||:::| |  |||||||||||   ||  |   |||  ||||||||:||||||||||
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
              |:|||||| |:|||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
              |||||||||||||||:|||:|||||: ||||||||||::||||::||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
              |:|||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 483
BD267475
LOCUS       BD267475                678 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION   BD267475
VERSION     BD267475.1  GI:33077243
KEYWORDS    JP 2002534980-A/16.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Baccala,G.P., Mallet,F. and Voisset,C.
  TITLE     Endogenetic nucleic fragment associated with an autoimmune disease,
            marking method and reagent
  JOURNAL   Patent: JP 2002534980-A 16 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534980-A/16
            PD   22-OCT-2002
            PF   21-JAN-2000 JP 2000594929
            PR   21-JAN-1999 FR     99/00888
            PI   GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC   G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenetic nucleic fragment associated with an autoimmune  CC
                 disease,
            CC   marking method and reagent
            FH   Key             Location/Qualifiers
            FT   source          1. .678
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| ||||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:||||||||||:|||||||||||||:||
Db         121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||              ||||||||||||
Db         181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:|||||||||||||||||:|||||||||||||||:::||||||||||||||||||
Db         241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db         301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                          |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db         361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db         420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||   ||  |   |||  ||||||||:||||||||||
Db         480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
```

```
Db         538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||||:|||:|||||: |||||||||||::||||::||||||:|||||||
Db         597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
             |:|||||||||||||||||||||
Db         657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 484
AX000984
LOCUS       AX000984                678 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 29 from Patent WO9902696.
ACCESSION   AX000984
VERSION     AX000984.1  GI:7241226
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Beseme,F. and Blond,J.
  TITLE     ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
            DISEASES OR WITH PREGNANCY DISORDERS
  JOURNAL   Patent: WO 9902696-A 29 21-JAN-1999;
            BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| ||||||||||||||||:||:|||::|||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             |||||||||||||||||||||||| :|||||:||||||||||:|||||||||||||:||
Db         121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||||              ||||||||||||
Db         181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:|||||||||||||||||:|||||||||||||||:::||||||||||||||||||
Db         241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db         301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                          |||||| |||:||||||| |||| ||: |||:|||||||||| :||
Db         361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:|||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db         420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||   ||  |   |||  ||||||||:||||||||||
Db         480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537
```

```
Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||||:|||:|||||: |||||||||||::||||::||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
             |:||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 485
AX027468
LOCUS       AX027468                678 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 18 from Patent FR2788784.
ACCESSION   AX027468
VERSION     AX027468.1  GI:10188432
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mallet,F., Voisset,C. and Paranhos,B.G.
  JOURNAL   Patent: FR 2788784-A 18 28-JUL-2000;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| ||||||||||||||:||:|||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||| :|||||:||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||              ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
             |:|| |:||||||||||||||||:||||||||||||||:::|||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
             |||||||| |||:|||||||||||| |||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:|||||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
             |||:|||||:||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
             ||||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||
```

```
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
             |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
             |||||||||||||||:|||:|||||: |||||||||||::||||::||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
             |:||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 486
AF072494
LOCUS       AF072494                678 bp    DNA     linear   PRI 10-FEB-1999
DEFINITION  Homo sapiens multiple sclerosis associated retrovirus pol probe
            Ppol-MSRV sequence.
ACCESSION   AF072494
VERSION     AF072494.1  GI:4262278
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
            Mandrand,B. and Mallet,F.
  TITLE     Molecular characterization and placental expression of HERV-W, a
            new human endogenous retrovirus family
  JOURNAL   J. Virol. 73 (2), 1175-1185 (1999)
   PUBMED   9882319
REFERENCE   2  (bases 1 to 678)
  AUTHORS   Blond,J.-L., Beseme,F. and Mallet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1998) UM103 CNRS-bioMerieux, ENS Lyon, 46 allee
            d'Italie, Lyon, Cedex 07 69364, France
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="isolated from synoviocyte cultures of a rheumatoid
                     arthritis (RA) patient"
     misc_feature    1. .678
                     /note="MSRV pol probe Ppol-MSRV"
     repeat_region   1. .678
                     /rpt_family="multiple sclerosis associated retrovirus"

  Query Match             6.7%;  Score 505.2;  DB 1;  Length 678;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  553;  Conservative   63;  Mismatches   32;  Indels   34;  Gaps    5;

Qy       2908 TCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCAT 2967
             ||||||||||:|||||||||||||||||||||||||||||||||||||||:||:|::||||
Db          1 TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGCCAGTTCTCAT 60

Qy       2968 ACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAA 3027
             ||||||| | |||||||||| :|||:| ||||||||||||||:||:|||::||||||||
Db         61 ACCTGGATATTCTTGTCCTTTGGTATGCGGATGATTTACTTTTAGCCGCCCGTTCAGAAA 120

Qy       3028 CCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGG 3087
             ||||||||||||||||||||||| :|||||:||||||||||:|||||||||||||:||
Db        121 CCTTGTGCCATCAAGCCACCCAAGTGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGG 180

Qy       3088 TTTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAG 3135
             |||||||||:|:|:|||||:|||||||||||||              ||||||||||
Db        181 TTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGAAGGCTATTTACCCTAAATACTTAG 240

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
```

```
                |:|| |:|||||||||||||||||:|||||||||||||||||:::|||||||||||||||||
Db        241 GGCTGAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 300

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGY---- 3251
                |||||||| |||:|||||||||||| ||||||||: ::|||||||||:|||:|||:
Db        301 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 360

Qy       3252 --------------TTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                           |||||| |||:|||||| |||| ||: |||:|||||||||| :||
Db        361 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 419

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
                |||:|||||:|||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        420 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 479

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
                ||||:::| |  |||||||||||    ||  |   |||  |||||||:|||||||||
Db        480 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 537

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
                |:|||||| |:|||||||||||||:||: | ::|||||| |||||| ||:||:|||| |:
Db        538 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTGTGTCACAG-AAAAAATAGGAATAGCTNTA 596

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
                |||||||||||||||:|||:||||||: ||||||||||::||||::||||||:|||||||
Db        597 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 656

Qy       3538 AYTGATGTAGTGGCAAAGGGTT 3559
                |:||||||||||||||||||||
Db        657 ATTGATGTAGTGGCAAAGGGTT 678


RESULT 487
AB114918
LOCUS       AB114918                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-3.
ACCESSION   AB114918
VERSION     AB114918.1  GI:41016426
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-3"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 505.1;  DB 1;  Length 537;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  522;  Conservative    3;  Mismatches   12;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                |||||||  |||||||||||||||||||||||||||||||||||  ||||   ||||||||

Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCAGTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                | | |||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 488
AB051010
LOCUS       AB051010                540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-X-3.
ACCESSION   AB051010
VERSION     AB051010.1  GI:14020839
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-X-3"
                     /db_xref="taxon:9606"
                     /chromosome="X"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
```

```
                     /product="envelope protein"
                     /protein_id="BAB47562.1"
                     /db_xref="GI:14020840"
                     /translation="NNFSTEINTTSVLVGPLVSNLEITHTSNLTCVKFSNTIDTTSSQ
                     CIRWVTPPTRIVCLPSGIFFVCGTSAYHCLNGSSESMCFLSFLVPPMTIYTEQDLYNH
                     VVPKPHNKRVPILPFVIRAGVLGRLGTGIGGITTSTQFYYKLSQEINGDMEQVTDSLV
                     TLQDQLNSLAAVVLQNRRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             6.7%;  Score 504.8;  DB 1;  Length 540;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    4;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || |||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 489
AB115112
LOCUS       AB115112                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE5-1.
ACCESSION   AB115112
VERSION     AB115112.1  GI:41016547
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE5-1"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 504.8;  DB 1;  Length 540;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    4;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||||||||||| ||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||| ||||||||||||| ||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 490
AB115120
LOCUS       AB115120                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE6-5.
ACCESSION   AB115120
VERSION     AB115120.1  GI:41016555
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
```

```
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE6-5"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 504.8;  DB 1;  Length 540;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    4;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||| |||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 491
AB120044
LOCUS       AB120044                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG1-1.
ACCESSION   AB120044
VERSION     AB120044.1  GI:45330715
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG1-1"
                     /isolation_source="brain of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 504.3;  DB 1;  Length 520;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  510;  Conservative    9;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||||:|||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGGTTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              |||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 492
AB120048
LOCUS       AB120048                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG1-5.
ACCESSION   AB120048
VERSION     AB120048.1  GI:45330719
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
```

```
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG1-5"
                     /isolation_source="brain of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 504.3;  DB 1;  Length 520;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  510;  Conservative    9;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||:||||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||  ||||||||||||||||||||||||||||  ||||||||||||||||||||
Db        361 TGAACCAAAGGGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::||||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 493
AB120049
LOCUS       AB120049                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG3-1.
ACCESSION   AB120049
VERSION     AB120049.1  GI:45330720
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG3-1"
                     /isolation_source="testis of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 504.3;  DB 1;  Length 520;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  510;  Conservative    9;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||| ||||||||||||||||:||||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCGTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::||||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 494
AB120057
LOCUS       AB120057                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG10-2.
ACCESSION   AB120057
VERSION     AB120057.1  GI:45330728
KEYWORDS    .
```

```
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG10-2"
                     /isolation_source="spleen of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 504.3;  DB 1;  Length 520;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches  510;  Conservative    9;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||  ||||||||||||||||:||||||||:||||||||||||||||||||||||
Db          1 AAACGCCCTTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 495
AB115116
LOCUS       AB115116                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE5-5.
ACCESSION   AB115116
VERSION     AB115116.1  GI:41016551
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE5-5"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.5;  DB 1;  Length 537;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  521;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||| |||||||||| ||||  ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 496
AB115124
```

```
LOCUS       AB115124                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE7-5.
ACCESSION   AB115124
VERSION     AB115124.1  GI:41016559
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE7-5"
                     /db_xref="taxon:9606"
                     /cell_line="Jurkat"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.5;  DB 1;  Length 537;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  521;  Conservative    3;  Mismatches   13;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 497
AB115103
LOCUS       AB115103                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE2-5.
ACCESSION   AB115103
VERSION     AB115103.1  GI:41016538
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE2-5"
                     /db_xref="taxon:9606"
                     /cell_line="PFSK-1"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.2;  DB 1;  Length 540;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  ||||||||||||| ||| ||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||  |||||||   | |||||||||||||| ||| ||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480
```

```
Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||| ||||||| |||||||||||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540

RESULT 498
AB115132
LOCUS       AB115132                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE9-2.
ACCESSION   AB115132
VERSION     AB115132.1  GI:41016567
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE9-2"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.2;  DB 1;  Length 540;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || ||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||   |||||||||||||||||||| ||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540

RESULT 499
AB115150
LOCUS       AB115150                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE13-2.
ACCESSION   AB115150
VERSION     AB115150.1  GI:41016585
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE13-2"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.2;  DB 1;  Length 540;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || ||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360
```

```
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 500
AB115133
LOCUS       AB115133                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE9-3.
ACCESSION   AB115133
VERSION     AB115133.1  GI:41016568
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE9-3"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 503.1;  DB 1;  Length 537;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  520;  Conservative    4;  Mismatches   13;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||| |||||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        118 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||||| ||||||||||||||||||||||||||||||||||||
Db        298 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 501
AB115303
LOCUS       AB115303                686 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP17-4.
ACCESSION   AB115303
VERSION     AB115303.1  GI:41052531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 686)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP17-4"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
                     /note="HERV-W pol"
     repeat_region   1. .686
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 503.1;  DB 1;  Length 686;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  558;  Conservative   58;  Mismatches   37;  Indels   35;  Gaps    5;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGGR 3137
             | || | |: :|||||:|| ||||||||||||               ||||||||||||:
Db          3 CAAACCAAAGGGCTCAGCTTTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGGG 62

Qy       3138 CTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCT 3197
             ||||:||||||||||||||:|||||||||||||||:::|||||||||||||||||||||||
Db         63 CTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCT 122

Qy       3198 TATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT------------------- 3239
             |||||| |||:||||||||||| ||||||||: ::||||||
Db        123 TATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAACA 182

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
             |||:|||:|||:|||||| |||:|||||| |||| ||: |||:|||||||||| :||||
Db        183 GGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCATT 241
```

```
Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGA 3359
Db        242 ATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTGA 301

Qy       3360 AGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGY 3419
Db        302 AGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAGC 359

Qy       3420 TTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGG 3479
Db        360 TTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAGG 418

Qy       3480 AGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAY 3539
Db        419 AGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAAT 478

Qy       3540 TGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKT 3599
Db        479 TGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCTT 538

Qy       3600 AGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGA 3659
Db        539 AGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATGA 598

Qy       3660 KGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAA 3719
Db        599 TGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTAA 658

Qy       3720 ATRTCAGGCTCTATTACTTGAARGGCCA 3747
Db        659 ATATCAGGCTCTATTACTTGAAGGGCCA 686
```

```
RESULT 502
AB079502
LOCUS       AB079502                542 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-2.
ACCESSION   AB079502
VERSION     AB079502.1  GI:22122076
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 542)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-2"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .542
                     /rpt_family="HERV-W"

  Query Match             6.7%;  Score 503;  DB 1;  Length 542;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches  527;  Conservative    4;  Mismatches    9;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        361 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACAT-GGAACGGGTCGCCGACTCCCTGGT 6677
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGGAACGGGTCGCCGACTCCTGGTC 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGT-CCTTCRAAATCGAAGAGCTT 6736
Db        481 ACCTTGGAAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTCAAAATCGAAGAGCTT 540

Qy       6737 TA 6738
Db        541 TA 542
```

```
RESULT 503
AB115213
LOCUS       AB115213                684 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP7-6.
ACCESSION   AB115213
VERSION     AB115213.1  GI:41016643
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 684)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .684
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP7-6"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
                     /note="HERV-W pol"
```

```
     repeat_region   1. .684
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 503;  DB 1;  Length 684;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches  558;  Conservative   58;  Mismatches   38;  Indels   34;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGGR 3137
              |||||||:|:|:| |||:|| || || ||||||              | || | |||||:
Db          1 TCCAAACCAAAGGGTCAGCTTTGTTCCCAGCAGAGGGCTATTTATCCTTAAAAATTAGGG 60

Qy       3138 CTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCT 3197
              ||||:||||||||||||||||:||||||||||||||:::||||||||||||||||||||
Db         61 CTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGCT 120

Qy       3198 TATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT------------------ 3239
              |||||| |||:|||| || |||| ||||||||: ::||||||
Db        121 TATCCTTATCCCAAACCCTTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAACA 180

Qy       3240 GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATT 3299
              |||:|||:|||:|||||| |||:||||||| |||| ||: |||:|||||||||| :||||
Db        181 GGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCATT 239

Qy       3300 AWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGA 3359
              |:|||||:|||||||||||||||||||||||||||||||| |||||||||:|:||||
Db        240 ATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTATTAAGATGGACACCTGA 299

Qy       3360 AGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGY 3419
              ||:::||||||||||||||||  ||  | | |||  ||||||||:|||||||||||||:
Db        300 AGCAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGC 357

Qy       3420 TTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGG 3479
              ||||||||:|||||||| ||||:||:|||::|||||| |||||||||:||:||||||:|
Db        358 TTGCCAACGGGGCAAGAGTTTTCTTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGT 416

Qy       3480 AGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAY 3539
              ||||||||||||:|| :|||||:||||||||||:: |||::||||||:|||||  ||:
Db        417 AGTCCTTACACAGGTCTGAGGGATGAGCTTGCAACCCACGGCATACCTGAGTAAGGCAAT 476

Qy       3540 TGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKT 3599
              ||||||  ||||||||||||||:|:||||||||||:|||||||| ||||||||||||::|
Db        477 CGATGTAATGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCTT 536

Qy       3600 AGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGA 3659
              |||||||||||||||||||||||||||:||||||||||||||||||| ||||||||
Db        537 AGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACTTCTCATGA 596

Qy       3660 KGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAA 3719
              :|||||: |||||||||||||||||||||||||||||||||||||||: |||| ||||
Db        597 TGTGAACTGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCATTTATTTAA 656

Qy       3720 ATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||:||||||||||||||||||:||| |
Db        657 ATATCAGGCTCTATTACTTGAAGGGCTA 684


RESULT 504
AB120050
LOCUS       AB120050                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG3-3.
ACCESSION   AB120050
VERSION     AB120050.1  GI:45330721
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG3-3"
                     /isolation_source="testis of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 502.7;  DB 1;  Length 520;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches  509;  Conservative    9;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||| ||||||||||||||:||||||||:|||||||||||||||||||||||||
Db          1 AAACGCCCTTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              |||||||||||||||||||||||||||:||||::||||||||| ||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGGCTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 505
AB120053
LOCUS       AB120053                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG8-1.
ACCESSION   AB120053
VERSION     AB120053.1  GI:45330724
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
```

```
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG8-1"
                     /isolation_source="placenta of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 502.7;  DB 1;  Length 520;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches  509;  Conservative    9;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||| |||||||||| ||||:||||||||:|||||||||||||||||||||||||
Db          1 AAACGCCCTTAAGACGTATTTTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 506
AB120055
LOCUS       AB120055                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG8-4.
ACCESSION   AB120055
VERSION     AB120055.1  GI:45330726
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG8-4"
                     /isolation_source="placenta of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 502.7;  DB 1;  Length 520;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches  509;  Conservative    9;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||||||||| ||||||||||:||||||||:|||||||||||||||| ||||||
Db          1 AAACGCCCCTAAGGCGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACAATAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              ||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 507
AB115224
LOCUS       AB115224                 685 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP9-5.
ACCESSION   AB115224
VERSION     AB115224.1  GI:41016654
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
          Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 685)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP9-5"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
                     /note="HERV-W pol"
     repeat_region   1. .685
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 502.4;  DB 1;  Length 685;
  Best Local Similarity   81.0%;  Pred. No. 0;
  Matches  558;  Conservative   58;  Mismatches   39;  Indels   34;  Gaps    5;

Qy       3089 TTCCAAACSARARGCTCARCTCTGCTCACAGCAG------------GTTAAATACTTAGG 3136
                | ||||  | |: :| |||:|| ||||||||||||           -|||||||||||
Db          1 TCCCAACCAAAGGGTTCAGCTTTGCTCACAGCAGAGGGCTATTTATCCTAAATACTTAGG 60

Qy       3137 RCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGC 3196
              :||||:|||||||||||||| |||:||||||||||||||:::||||||||||||||||||
Db         61 GCTAAAATTATCCAAAGGCCCCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGGC 120

Qy       3197 TTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT----------------- 3239
              ||||||| |||:|||||||||||| ||||||||: ::||||||
Db        121 TTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAAC 180

Qy       3240 -GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCAT 3298
               |||:|||:|||:|||||| |||:||||||| |||| ||: |||:|||||||||| :|||
Db        181 AGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCAT 239

Qy       3299 TAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTG 3358
              ||:|||||:|||||||||||||||||||||||||||||||||:|||||||||:|:|||
Db        240 TATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCTG 299

Qy       3359 AAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAG 3418
              |||:::| |  |||||||||||||   ||  |   |||  ||||||||:||||||||||||
Db        300 AAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAAG 357

Qy       3419 YTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRG 3478
              :|||||| |:|||||||||||||:||: ||::|||||| |||||| ||:||:|||||||:|
Db        358 CTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTAG 416

Qy       3479 GAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAA 3538
              |||||||||||||:|||:|||||: ||||||||||::|||||:|||||||:||||||||
Db        417 GAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAAA 476

Qy       3539 YTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYK 3598
              :|||||||||||||||||||:|:||||||||||:||||||||| ||||||||||||::
Db        477 TTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTCT 536

Qy       3599 TAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
              |||||||||||||||||||||||||| ||:|||||||||||||||| || |||||||
Db        537 TAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCATG 596

Qy       3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTA 3718
              |:|||||: |||||||||||||||||| ||||| |||||||||||||||:|||| ||||
Db        597 ATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTTA 656
```

```
Qy       3719 AATRTCAGGCTCTATTACTTGAARGGCCA 3747
              |||:||||||||||||||||||:|||||
Db        657 AATATCAGGCTCTATTACTTGAAGGGCCA 685


RESULT 508
AB126103
LOCUS       AB126103                660 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP1-10.
ACCESSION   AB126103
VERSION     AB126103.1  GI:47169703
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 660)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP1-10"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W pol"
     repeat_region   1. .660
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 502.2;  DB 1;  Length 660;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  558;  Conservative   57;  Mismatches   35;  Indels   18;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||||||||||||||||| |||:||||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAATACTCAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:|||||||||||||:::|||||||| |||||||||||||||||||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||||: : |||||||||:|||:|||:|||||||||: ||||||
Db        121 CAAACCCTAAAGCAACTAAGAAGCTTCCTTGGCATAACAGGCTTCTGCCGAATCTGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||:  | :||| |||||||:|||||:||||   ||||||| ||||||||||||
Db        181 -CCCAGGTACAGTGAAAGAGCCAGGCCATTATATAC---AATTAAGAAAACTCAGAAAGT 236

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              ||||||||||||:|||||||||:| ||||||:::|||||||||||||||||  ||  |
Db        237 CAATACCCATTTAGTAAGATGGACATCTGAAGCCAAAGTGGCTTTCCAGGCCCTAAAGAA 296

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  ||||||||:||||||||||||:||||||||: |||||||| |||:||:| |:
Db        297 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACGAGGCAAGAC-TTTCTTTAAAT 353

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:|||||| :|||||||||||||| :|||:|||||: |||||
Db        354 GTCACAG-AAAAAACAGGAATAGCTCCAGGAGTCCTTACACACATCCGAGGGATCAGCTT 412

Qy       3510 GCAACCYR----------TGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTT 3559
```

```
                |  ||||::          ||||::||||||:||| ||||:||||||||||||||||||||
Db        413 GTAACCTGAGTAATCAGCTGGCATACCTGAGTAAAGAAATTGATGTAGTGGCAAAGGGTT 472

Qy       3560 GRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAA 3619
                |:|:||||||||||:|||||||||   ||||||||||::||||||||||||||||||||
Db        473 GGCCTCATTGTTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAA 532

Qy       3620 TAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTG 3679
                ||||||||:|||||||||||||||||||| ||||||||||:|||||::||||||||||||
Db        533 CGATACAGGGAAGAGATCTTACTGTGTGGGCATCTCATGATGTGAATGGCATACTCACTG 592

Qy       3680 CTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTG 3739
                |||||||||||||||||||||||||||| :|||| ||||||||:|||||||||||||||||
Db        593 CTAAAGGAGACTTGTGGCTGTCAGACAATCGTTTGCTTAAATATCAGGCTCTATTACTTG 652

Qy       3740 AARGGCCA 3747
                ||:| |||
Db        653 AAGGACCA 660


RESULT 509
AB115284
LOCUS       AB115284                652 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP9-2.
ACCESSION   AB115284
VERSION     AB115284.1  GI:41052512
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 652)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP9-2"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W pol"
     repeat_region   1. .652
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 502.1;  DB 1;  Length 652;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  555;  Conservative   57;  Mismatches   36;  Indels   11;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
                |||||||:|:|:|||||:|||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
                ||||||||||:||| |||||||||||:::|||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
                |||||||||||||||||| :|:: ||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
                |||||||: |||:|||||||||  |||||:|     ||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
                |||||||||||||:|||||||||:|:|||| |:::|||  ||||||||||| |  ||  |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
                | |||  ||||||||:|||| ||||||:||||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
                :|||||| |||||||||:||:|||||||||||||||||:|||:|||||:||||||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
                ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA-GG 3628
                ||||:||||| ||| | ||||||||||::|||||||||||  |||||||||| ||||| ||
Db        472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
                |:|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        532 GAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 591

Qy       3689 ACTTGTGGC-TGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
                |||||||| ||||||||||:| ||||||||||:|||||||||||||||  |:|| |
Db        592 ACTTGTGGCTTGTCAGACAACTGGTTACTTAAATATCAGGCTCTATTACTTTGAAGGTC 650


RESULT 510
AB055412
LOCUS       AB055412                540 bp    DNA     linear   PRI 07-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:A549-2.
ACCESSION   AB055412
VERSION     AB055412.1  GI:15128494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Lee,W.-H., Kim,H.-M. and Kim,H.-S.
  TITLE     Identification of new endogenous retroviral sequences belonging to
            the HERV-W family in human cancer cells
  JOURNAL   Intervirology (2001) In press
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="A549-2"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.7%;  Score 502;  DB 1;  Length 540;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    3;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
```

```
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCTCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540
```

RESULT 511
AB115121
LOCUS AB115121 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:CWE6-6.
ACCESSION AB115121
VERSION AB115121.1 GI:41016556
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWE6-6"
/db_xref="taxon:9606"
/cell_line="UO-31"
/note="HERV-W env"
repeat_region 1. .540
/note="HERV-W env"

Query Match 6.7%; Score 502; DB 1; Length 540;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 514; Conservative 3; Mismatches 23; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540
```

RESULT 512
AB115119
LOCUS AB115119 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:CWE6-4.
ACCESSION AB115119
VERSION AB115119.1 GI:41016554
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWE6-4"
/db_xref="taxon:9606"
/cell_line="UO-31"
/note="HERV-W env"

```
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 501.6;  DB 1;  Length 540;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    4;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||| ||||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| |||||||| |||| ||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 513
AB115122
LOCUS       AB115122                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE7-1.
ACCESSION   AB115122
VERSION     AB115122.1  GI:41016557
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE7-1"
                     /db_xref="taxon:9606"
                     /cell_line="Jurkat"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 501.6;  DB 1;  Length 540;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    4;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||| |||
Db        301 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  ||||||| |||||||||||||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||| ||||| |||||||| |||| ||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 514
AB120046
LOCUS       AB120046                 520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG1-3.
ACCESSION   AB120046
VERSION     AB120046.1  GI:45330717
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
```

```
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG1-3"
                     /isolation_source="brain of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.7%;  Score 501.1;  DB 1;  Length 520;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  508;  Conservative    9;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              ||||||||||||||||||||||||:||||||||:||||||||||||||||||||||||||
Db          1 AAACGCCCCTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              | ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 AGCACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCTCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              |||||||||||||||||||||||||:||||::|| ||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCCCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 515
AB115144
LOCUS       AB115144                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE11-5.
ACCESSION   AB115144
VERSION     AB115144.1  GI:41016579
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES    Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE11-5"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 500.4;  DB 1;  Length 540;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    3;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||| ||||| |||||||||| |||| ||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 516
AB115168
LOCUS       AB115168                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE16-4.
ACCESSION   AB115168
VERSION     AB115168.1  GI:41016603
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
```

```
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE16-4"
                     /db_xref="taxon:9606"
                     /cell_line="2F7"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 500.4;  DB 1;  Length 540;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    3;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||| ||||||| ||  ||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||||||||||||||:||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACCCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 517
AB126125
LOCUS       AB126125                640 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W pol, isolate:JM-WP7-8.
ACCESSION   AB126125
VERSION     AB126125.1  GI:47169725
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 640)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WP7-8"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W pol"
     repeat_region   1. .640
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 500.1;  DB 1;  Length 640;
  Best Local Similarity   82.8%;  Pred. No. 0;
  Matches  545;  Conservative   59;  Mismatches   33;  Indels   21;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||| ||||||||||||||||||||:||||:|||||||
Db          1 TCCAAACCAGAGGCTCAGCTCTGGTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||||||||||||||:::|||||||| |||||||||||||| |||:|
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACGTATCCAGCCCATACTGGCTTATCCTAATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||  |||||||:|:: |||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAAGAACTAAGAGGGCTCCTTGGCATAACAGGCTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||: ||||||||||:||||||:||||||:|||||||
Db        181 -CCCAGGTACGGCAGAATAGCCAGGCCATTATATACACTAATTAA---------------- 224

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             ||||||||||||:||||||||:|:||||||:::|||||||||||||||||||  ||  |
Db        225 -AATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTCCAGGCCCTAAAGAA 283

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||||||||| :|||||| |:||||| |||||||:||:|||:
Db        284 GGCCCT--AACCCAAGCCCCAGTGTTAAACTTGCCAGCGGGGCAGGACTTTTCTTTATAT 341

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:||||||:|||||||||||||||:|| :|||||:||||||
Db        342 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCTGAGGGATGAGCTT 400

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::|||| :|||| |:||||||||:|||||||||||||| ||||||:|:||||||
Db        401 GCAACCCGTGGCCTACCTAAGTAAGGAAATTGATGTAGTGGCAAAAGGTTGGCCTCATTG 460

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||: |||||||| |||||||||| ||::|||||||||||||||||||||||| | ||
Db        461 TTTACAGGTAGTGGCGGCAGTAGCGGTCTTAGTATCTGAAGCAGTTAAAATAATATAAGG 520

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||| |||||||:|||||::|||| ||  |||||||| ||||
Db        521 AAGAGATCTTACTGTGTGGACAGCTCATGATGTGAACAGCATTCTTTCTGCTAAAAGAGA 580

Qy       3690 CTTGTGGCTGTCAGACAACYGTTT-ACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||: ||| ||||||||:|||||||||||||||||:|| |
Db        581 CTTGTGGCTGTCAGACAACCATTTCACTTAAATATCAGGCTCTATTACTTGAAAGGGC 638


RESULT 518
AB115006
LOCUS       AB115006                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-8.
```

```
ACCESSION   AB115006
VERSION     AB115006.1  GI:41016514
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-8"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 500;  DB 1;  Length 540;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    4;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||| |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db          1 AACAACCTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || ||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540

RESULT 519
```

```
AB115306
LOCUS       AB115306                686 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP18-2.
ACCESSION   AB115306
VERSION     AB115306.1  GI:41052534
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 686)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP18-2"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W pol"
     repeat_region   1. .686
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 500;  DB 1;  Length 686;
  Best Local Similarity   80.6%;  Pred. No. 0;
  Matches  556;  Conservative   59;  Mismatches   39;  Indels   36;  Gaps    5;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAG-------------GTTAAATACTTAG 3135
             |||||||:|:|:|||||:|| |||||||||||              | ||    | |
Db          1 TCCAAACCAAAGGCTCAGCTTTGCTCACAGCAGAGGGCTATTTATCCCTAAATACTTAGG 60

Qy       3136 GRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGG 3195
              |:||||:|||||||||||||||:|||||||||||||||:::|||||||||||||||
Db         61 GGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGGAATGTATCCAGCCTATACTGG 120

Qy       3196 CTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTT---------------- 3239
              |||||||| |||:|||||||||||| ||||||||: ::||||||
Db        121 CTTATCCTTATCCCAAAACCCTAAAACAACTAAGAAGGTTCCTTGGCATAATAGGCATAA 180

Qy       3240 --GGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCA 3297
                |||:|||:|||:|||||| ||||:|||||||| |||| ||: |||:|||||||||| :||
Db        181 CAGGCATAACAGGTTTCTGCTGAATATGGATT-CCCAAGTACGGCAAAATAGCCAGACCA 239

Qy       3298 TTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCT 3357
              |||:||||||||||||||||||||||||||||||||||||||||:|||||||||:|:||
Db        240 TTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACACCT 299

Qy       3358 GAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAA 3417
              ||||:::| |  |||||||||||   ||  |  |||  ||||||||:||||||||||
Db        300 GAAGCAGAGGCAGCTTTCCAGGCCGTAAAGAACACCCT--AACCCAAGCCCCAGTGTTAA 357

Qy       3418 GYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTR 3477
              |:|||||| |:||||||||||||:||: ||::|||||| |||||| ||:||:|||||:
Db        358 GCTTGCCAGCGGGGCAAGACTTTTCTTTCTATGTCACAG-AAAAAATAGGAATAGCTCTA 416

Qy       3478 GGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAA 3537
              |||||||||||||:|||:|||||: |||||||||||::||||::||||||:|||||||
Db        417 GGAGTCCTTACACAGGTCCGAGGGACCAGCTTGCAACCCATGGCATACCTGAGTAAGGAA 476

Qy       3538 AYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTY 3597
              |:|||||||||||||||||||:|:||||||||||:||||||||| ||||||||||||:
```

```
Db        477 ATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTACGGGTAGTGGCGGCAGTAGCAGTC 536

Qy       3598 KTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCAT 3657
              :|||||||||||||||||||||||||||| ||:|||||||||||||||||| || ||||||
Db        537 TTAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAGATCTTACTGTGTGAACCTCTCAT 596

Qy       3658 GAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTT 3717
              ||:||||||: |||||||||||||||||| ||||| |||||||||||||||||:|||| |||
Db        597 GATGTGAACCGCATACTCACTGCTAAAGAAGACTGGTGGCTGTCAGACAACTGTTTGCTT 656

Qy       3718 AAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||:|||||||||||||||||||:|||||
Db        657 AAATATCAGGCTCTATTACTTGAAGGGCCA 686


RESULT 520
AB115301
LOCUS       AB115301                645 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP15-8.
ACCESSION   AB115301
VERSION     AB115301.1  GI:41052529
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 645)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP15-8"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W pol"
     repeat_region   1. .645
                     /note="HERV-W pol"

  Query Match             6.7%;  Score 499.5;  DB 1;  Length 645;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  551;  Conservative   60;  Mismatches   30;  Indels   19;  Gaps    6;

Qy       3091 CCAAACSARARGCTCARCTCTGCTCACAGCAGGTT-AAATACTTAGGRCT-AARATTATC 3148
              ||||||:|:|:| |||:||||||||||||||||||| ||||||||||:|| ||:||||||
Db          2 CCAAACCAAAGGGTCAGCTCTGCTCACAGCAGGTTAAAATACTTAGGGCTAAAAATTATC 61

Qy       3149 CAAAGGCACCA-RGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATC 3207
              |||||||||||| :|||||||||||||||||::: |||||||||||||||||||||||||||
Db         62 CAAAGGCACCAGGGGCCCTCAGTGAGGAACGTTTCCAGCCTATACTGGCTTATCCTCATC 121

Qy       3208 YCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGA 3267
              :||||||||||||||||||||:|::||||||||||:|||:|||:||||||||||:|||||
Db        122 CCAAAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCGAATATGGA 181

Qy       3268 TTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAA 3327
              || |||||| :  | :||||||||||:|||||:|||||:|||||| |||||||||| |
Db        182 TT-CCCAGGCACAGTGAAATAGCCAGGCCATTATATATACACTAATTATGGAAACTCAGACA 240

Qy       3328 GCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAA 3387
              || ||||||||||:|||||||||:|:||||||:::||| |||||||||||||
Db        241 ACCTATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCC----- 295

Qy       3388 GAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYAT 3447
                          ||||||||:|| ||  |||||:||||||| :|||||||||||||:||:||
Db        296 ---------TAACCCAAGCCCTAGCATTAAGCTTGCCAATGGGGCAAGACTTTTCTTTAT 346

Qy       3448 AYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGC 3507
              |::||||||  |||||||||:||:||||||:|||||||||||||||:|||:|||||: |||
Db        347 ATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACCAGC 405

Qy       3508 TTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCAT 3567
              ||||||||::||||::||||||:||||||||:|||||  |||||||||| ||||:|:|||
Db        406 TTGCAACCCATGGCACACCTGAGTAAGGAAATTGATGGAGTGGCAAAGAGTTGGCCTCAC 465

Qy       3568 TGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAG 3627
              ||||| :|||||||||  |||||||||| |::||||||| ||||||||||||||||||||
Db        466 TGTTTTTGGGTAGTGGCAGCAGTAGCAATCTTAGTATCCGAAGCAGTTAAAATAATACAG 525

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              ||:|||||||||||||||||||||||||||||||:|| ||::  ||||||||||||||||
Db        526 GGAAGAGATCTTACTGTGTGGACATCTCATGATGTAAATGATATACTCACTGCTAAAGGA 585

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||||||||||||||||||: ||| |||||||:||||||||||||||||:|||||
Db        586 GACTTGTGGCTGTCAGACAACCATTTGCTTAAATATCAGGCTCTATTACTTGAAGGGCCA 645


RESULT 521
AB114987
LOCUS       AB114987                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-12.
ACCESSION   AB114987
VERSION     AB114987.1  GI:41016495
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-12"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.7%;  Score 499.1;  DB 1;  Length 539;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches  518;  Conservative    3;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAGCAATACTACATACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        420 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539
```

RESULT 522
AB115102
LOCUS AB115102 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:CWE2-4.
ACCESSION AB115102
VERSION AB115102.1 GI:41016537
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWE2-4"
/db_xref="taxon:9606"
/cell_line="PFSK-1"
/note="HERV-W env"
repeat_region 1. .540
/note="HERV-W env"

Query Match 6.7%; Score 498.8; DB 1; Length 540;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 512; Conservative 3; Mismatches 25; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||||||||||||||||||||||| |||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||| |||||||||  ||| ||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATAGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540
```

RESULT 523
AB115007
LOCUS AB115007 539 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE12-11.
ACCESSION AB115007
VERSION AB115007.1 GI:41016515
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 539)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE12-11"
/db_xref="taxon:9606"
/tissue_type="uterus"
repeat_region 1. .539
/note="HERV-W env"

Query Match 6.7%; Score 498.5; DB 1; Length 539;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 518; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||| ||||||| |||||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539
```

```
RESULT 524
AB051003
LOCUS       AB051003                 540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-6-3.
ACCESSION   AB051003
VERSION     AB051003.1  GI:14020828
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-6-3"
                     /db_xref="taxon:9606"
                     /chromosome="6"
     repeat_region   1. .540
                     /rpt_family="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACGCGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        301 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540
```

```
RESULT 525
AB114962
LOCUS       AB114962                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-3.
ACCESSION   AB114962
VERSION     AB114962.1  GI:41016470
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-3"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .540
```

```
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||| |||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||    | |||||||||||||| ||| ||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 526
AB114990
LOCUS       AB114990                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-1.
ACCESSION   AB114990
VERSION     AB114990.1  GI:41016498
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-1"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||| |||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCACATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||    | |||||||||||||| ||| ||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 527
AB115130
LOCUS       AB115130                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE8-5.
ACCESSION   AB115130
VERSION     AB115130.1  GI:41016565
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
```

```
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE8-5"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || |||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 528
AB115139
LOCUS       AB115139                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE10-5.
ACCESSION   AB115139
VERSION     AB115139.1  GI:41016574
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE10-5"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        121 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || | ||||||| |||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCCAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||   |||||||||||||||||| ||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 529
AB115143
LOCUS       AB115143                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE11-4.
ACCESSION   AB115143
VERSION     AB115143.1  GI:41016578
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
```

```
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE11-4"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.4;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||| || |||||||| ||||||||||||||||||
Db        121 CCCAACTCCCAATGCATCAGGTGGGTAAGTCATCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || ||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||   | ||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 530
AB114976
LOCUS       AB114976                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-8.
ACCESSION   AB114976
VERSION     AB114976.1  GI:41016484
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-8"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.3;  DB 1;  Length 537;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  517;  Conservative    4;  Mismatches   16;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | |||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 531
AB115153
LOCUS       AB115153                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE13-5.
ACCESSION   AB115153
VERSION     AB115153.1  GI:41016588
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE13-5"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.3;  DB 1;  Length 537;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  517;  Conservative    4;  Mismatches   16;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||   ||||||||||||||||||| ||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        418 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 532
AB114916
LOCUS       AB114916                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-1.
ACCESSION   AB114916
VERSION     AB114916.1  GI:41016424
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-1"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.7%;  Score 498.1;  DB 1;  Length 539;
  Best Local Similarity   95.7%;  Pred. No. 0;
  Matches  517;  Conservative    4;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAGTACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||  ||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||  || ||||||||||||||||||||||||||||||||||||||||||| |||
Db        300 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||  |||||||   | ||||||||||||| ||| ||||||||||||||||||||||||
Db        360 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||| ||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATTGAAGAGCTTTA 539


RESULT 533
AB120047
LOCUS       AB120047                520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG1-4.
```

```
ACCESSION   AB120047
VERSION     AB120047.1  GI:45330718
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG1-4"
                     /isolation_source="brain of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.6%;  Score 497.9;  DB 1;  Length 520;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  506;  Conservative    9;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||| |||||||||| ||||:||||||||:|||||||||||||||||||||||||
Db          1 AAACGCCCGTAAGACGTATTGTGGAGAATTGGGACCAATTTGACCCTCAGACACTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              |||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGGGTGATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              |||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||:|||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520


RESULT 534
AB115191
LOCUS       AB115191                649 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:WP1-7.
ACCESSION   AB115191
VERSION     AB115191.1  GI:41016621
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WP1-7"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
                     /note="HERV-W pol"
     repeat_region   1. .649
                     /note="HERV-W pol"

  Query Match             6.6%;  Score 497.3;  DB 1;  Length 649;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  552;  Conservative   57;  Mismatches   39;  Indels   11;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCA-GGTTAAATACTTAGGRCTAARATTATC 3148
              ||| |||: :|:| |||:||||| ||| |||| |||||||||| ||||:  ||:||||||
Db          1 TCCCAACCCAAGGGTCAGCTCTGGTCAAAGCAGGGTTAAATACGTAGGGGCAAAATTATC 60

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCY 3208
              ||||||| |||:||| |||||| |||||:::|||||||||||||||||||| |||||||:
Db         61 CAAAGGCCCCAGGGCACTCAGT-AGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCC 119

Qy       3209 CAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGAT 3268
              |||| |||||||||||||| :|:: ||||||||:|||:|||:||||||||:||||||
Db         120 CAAACCCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGAT 179

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
              | |||||||: |||:|||||||||||  |||||:|     ||||||||||||||||||||
Db         180 T-CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAG 233

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
              |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||| |  ||
Db         234 CCAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGA 293

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
              | | |||  ||||||||:||||| ||||||:||||||||:|| ||||||||||:|||
Db         294 AGGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATA 351

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
               :|||||| |||||||||:||:|||||:|||||||||||||||:|||:||||||:|||||
Db         352 AGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCT 410

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
              |||||||::||||::||||||:||||||||:||||||||||||||||  ||||:|: ||||
Db         411 TGCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATT 470

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
              |||||:||||| ||| | ||||||||||::|||||||||||  |||||||||  |||||||
Db         471 GTTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGG 530
```

```
Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
             |:|||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||
Db        531 GAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAG 590

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             ||||||||||||||||||||:||||||||||||:||||||||||||||||||:| |||
Db        591 ACTTGTGGCTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGTCCA 649


RESULT 535
AB115126
LOCUS       AB115126                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE8-1.
ACCESSION   AB115126
VERSION     AB115126.1  GI:41016561
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE8-1"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 497.2;  DB 1;  Length 540;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| ||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGCTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||| ||||:| | ||||||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTACGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 536
CQ742266
LOCUS       CQ742266                 627 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28200 from Patent WO02068579.
ACCESSION   CQ742266
VERSION     CQ742266.1  GI:42352438
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 28200 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             6.6%;  Score 497.1;  DB 1;  Length 627;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  510;  Conservative   38;  Mismatches   45;  Indels    1;  Gaps    1;

Qy       3779 ACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAA 3838
             ||  | :|  | |   | |  ||||||||||||||||:||:||||||||||||| :| |
Db         35 ACGTATCCAAAATATCTGCCTACAATGAAGAAAAGATAGAACATAACTGTCAACAGGTGA 94

Qy       3839 TTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACC 3898
             || |||||||||| ||  |||||||||||||:||||:|||||:||||||||||: |||
Db         95 TTGCTCAAACCTACGCTGCTCGAGGGGACCTTCTAGAGGTTCCCTTGACTGATCCCTACC 154

Qy       3899 TTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTAT 3958
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||:|||||||
Db        155 -TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTAT 213

Qy       3959 GCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCT 4018
             |||||||||||||||||||||||| :|||||||||||||||| ||||||||||||||  ||
Db        214 GCAGTGGTCAGTGATAATGGAATCCTTGAAAGTAATCCCCACACTCCAGGAACTAGCACT 273

Qy       4019 YAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYA 4078
              :| ||:|||||||||||||:|||:|  |||||||||||||||||| :|||||||:|
Db        274 CAACTGGCAGAACTAATAGCCCTCACTCAGGCACTAGAATTAGGAGAACGAAAAAGGGTA 333

Qy       4079 AATATATATACAGACTCTRARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGS 4138
             |||||||||||||||||:|:||||  :|||||||| |||||||||| |::|||||||:
Db        334 AATATATATACAGACTCTAAGTATGTCTACCTAGTCCTCCATGCCCACGCAGCAATATGG 393

Qy       4139 ARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTATCAMACATCAGGAAGCCATTAGG 4198
              |:|||||||||| |||||||||: ||:|||||||||||:|||||||||||||||||||||
Db        394 AGAGAAAGGGAATGCCTAACTTCCAAGGGAACACCTATCAAACATCAGGAAGCCATTAGG 453

Qy       4199 ARATTATTAYTGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGGTCAT 4258
              |:|||||||:||||:|||||||||||||:|||||||||:|||||||||||||:|||||||
Db        454 AGATTATTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACACTGCCGGGGTCAT 513
```

```
Qy       4259 CANAAAGGAAAGRAAAGGGAAATASAAGRGAAYTGCCAAGCAKATATTGAAGCMAAAAGA 4318
              || |||||||||:|||| |||||||:|||:|||:|||||||| :||||||||||:||||||
Db        514 CAGAAAGGAAAGGAAAGAGAAATAGAAGGGAACTGCCAAGCGGATATTGAAGCCAAAAGA 573

Qy       4319 GCTGCAAGGCAGGACCCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAG 4372
              || |||||||||||||||||||||||||||||| |  ||   ||| ||||||||
Db        574 GCCGCAAGGCAGGACCCTCCATTAGAAATGCTTTTAGAAGGACCCGTAGTATAG 627


RESULT 537
AB050998
LOCUS       AB050998                540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-3-7.
ACCESSION   AB050998
VERSION     AB050998.1  GI:14020821
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-3-7"
                     /db_xref="taxon:9606"
                     /chromosome="3"
     repeat_region   1. .540
                     /rpt_family="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTGGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTATGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||  ||||||||||||||  |||||||||||||||||| ||
Db        121 ACCAACTCCCAATGCATCAGGTGAATAACTCCTCCCACAGGAATAGTCTGCCTACCCCCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGCACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||  || ||||||| ||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTTGTACCTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||| |||||||||| |||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTGCTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||| |||||||||||||| |||||||| |||| ||||||||||||||||
Db        421 TACTACAAACTATCCCAAGAACTAAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||| ||||||||||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTGCAAGATCGACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 540


RESULT 538
AB051007
LOCUS       AB051007                540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-14-1.
ACCESSION   AB051007
VERSION     AB051007.1  GI:14020833
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-14-1"
                     /db_xref="taxon:9606"
                     /chromosome="14"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47559.1"
                     /db_xref="GI:14020834"
                     /translation="NNFSTEINTTSILVGPLVSNLEITHTSNLTRVKFSNTIDTTNSQ
                     CIRWVTPPTQIVCLPSGIFFVCGTSAYHCLNGFSESMCFLSFLVHPMTIYTEQDLYNY
                     VVPKPRNKRVPILPFVIRAGMLGRLGTGMGSITTSTQFYYKLSQELNGDMERVADSLV
                     TLQDQLNSLAAVVLQNRRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||    | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 539
AB114954
LOCUS       AB114954                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-6.
ACCESSION   AB114954
VERSION     AB114954.1  GI:41016462
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-6"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
```

```
             :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||    | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 540
AB115001
LOCUS       AB115001                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-2.
ACCESSION   AB115001
VERSION     AB115001.1  GI:41016509
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-2"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
```

```
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||| || |||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 541
AB115117
LOCUS       AB115117                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE6-1.
ACCESSION   AB115117
VERSION     AB115117.1  GI:41016552
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE6-1"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
```

```
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCGTGTAAAATTTAGCAATACTATAGACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||| || |||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540
```

```
RESULT 542
AB115123
LOCUS       AB115123                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE7-2.
ACCESSION   AB115123
VERSION     AB115123.1  GI:41016558
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE7-2"
                     /db_xref="taxon:9606"
                     /cell_line="Jurkat"
                     /note="HERV-W env"
     repeat_region   1. .540
```

```
                     /note="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||| ||| |||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| ||||||| | |||||||||||||| ||| ||||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 543
AB115142
LOCUS       AB115142                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE11-3.
ACCESSION   AB115142
VERSION     AB115142.1  GI:41016577
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE11-3"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 496.8;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| ||||||| | |||||||||||||| ||| ||||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 544
AB115160
LOCUS       AB115160                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE15-1.
ACCESSION   AB115160
VERSION     AB115160.1  GI:41016595
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
```

```
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE15-1"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 495.6;  DB 1;  Length 540;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||| |||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGCGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||| |||||||||| | ||||||| ||||||||| |||||||||||||||||||||| |
Db        361 GCAAGAGTGCTAGGCGGACTAGGTTCTGGCATTGGTGGTATCACAACCTCTACTCAGTCC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||||| |||||||||| ||| ||| ||||||| |||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACAGGTTGCCAACTCCCTCGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||| ||| ||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGCGCTTTA 540


RESULT 545
AB114982
LOCUS       AB114982                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-6.
ACCESSION   AB114982
VERSION     AB114982.1  GI:41016490
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-6"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 495.2;  DB 1;  Length 540;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || |||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||| ||||:||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 546
AB114996
LOCUS       AB114996                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-11.
ACCESSION   AB114996
VERSION     AB114996.1  GI:41016504
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
```

```
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-11"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 495.2;  DB 1;  Length 540;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||  |||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCACATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||| ||||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 540


RESULT 547
AB115141
LOCUS       AB115141                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE11-2.
ACCESSION   AB115141
VERSION     AB115141.1  GI:41016576
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE11-2"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 495.2;  DB 1;  Length 540;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||| ||||:|||| ||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATTGAAGAGCTTTA 540


RESULT 548
AB115288
LOCUS       AB115288                 652 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP10-2.
ACCESSION   AB115288
VERSION     AB115288.1  GI:41052516
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 652)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP10-2"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W pol"
     repeat_region   1. .652
                     /note="HERV-W pol"

  Query Match             6.6%;  Score 495.1;  DB 1;  Length 652;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  548;  Conservative   55;  Mismatches   34;  Indels   11;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
             |||||||:|:|:|||||:|||||||||| |||||||||||||| ||||:  ||:|||||||
Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
             ||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db         61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
             |||||||||||||||||| :|:: ||||||||:|||:|||:||||||||||:|||||||
Db        121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
             |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||||
Db        181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
             |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||| |  ||  |
Db        235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
             | |||  ||||||||:|||||  |||||:|||||||:|| ||||||||||:||:|||
Db        295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
             :|||||| |||||||||:||:|||||:|||||||||||||||:|||:|||||:||||||
Db        353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
             ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db        412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGG 3629
             ||||:||||| ||| | |||| |||||::|||||||||| |||||||| ||||||||
Db        472 TTTACGGGTATTGGCGACAGTCGCAGTCTTAGTATCTGAAAAAGTTAAAATGATACAGGG 531

Qy       3630 RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGA 3689
             :|||||||||||||||||||||||||||||:||||::||||||||||||||||||||||
Db        532 AAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGA 591

Qy       3690 CTT-GTGGCTGTCAGACAACYGTTTACTTAAA-TRTCAGGCTCTATTA 3735
             ||| ||||||||||||||:|||||||||||  | ||||||||||||
Db        592 CTTGGTGGCTGTCAGACAACTGTTTACTTAAATTTTCAGGCTCTATTA 639
```

```
RESULT 549
AB120052
LOCUS       AB120052                520 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG3-8.
ACCESSION   AB120052
VERSION     AB120052.1  GI:45330723
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 520)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG3-8"
                     /isolation_source="testis of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .520
                     /note="corresponds to gag"

  Query Match             6.6%;  Score 494.7;  DB 1;  Length 520;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  504;  Conservative    9;  Mismatches    7;  Indels    1;  Gaps    1;

Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACACTAAGAA 691
              |||||||| |||||||||| ||||:|||||||:| |||||||||||||| ||||||
Db          1 AAACGCCCGTAAGACGTATTGTGGAGAATTGGGACCCATTTGACCCTCAGACAGTAAGAA 60

Qy        692 AGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 751
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAGGGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTAT 120

Qy        752 AACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 811
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        121 AACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAA 180

Qy        812 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 871
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGC 240

Qy        872 CCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMY 931
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||::
Db        241 CCTACAGGAAGCCTTCAGAGTTTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAAC 300

Qy        932 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 991
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAG 360

Qy        992 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAG 1051
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        361 TGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAG 419

Qy       1052 AATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAA 1111
              |||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||
Db        420 AATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAA 479

Qy       1112 CAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||:||||||||||||||||
Db        480 CAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 520
```

```
RESULT 550
AB115110
LOCUS       AB115110                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE4-2.
ACCESSION   AB115110
VERSION     AB115110.1  GI:41016545
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE4-2"
                     /db_xref="taxon:9606"
                     /cell_line="HCT-116"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.9;  DB 1;  Length 537;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    3;  Mismatches   19;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||   |||||||||||||||||||| ||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 551
AB115154
LOCUS       AB115154                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE13-6.
ACCESSION   AB115154
VERSION     AB115154.1  GI:41016589
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE13-6"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.9;  DB 1;  Length 537;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    3;  Mismatches   19;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||   | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 552
AB114921
LOCUS       AB114921                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-8.
ACCESSION   AB114921
VERSION     AB114921.1  GI:41016429
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-8"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.7;  DB 1;  Length 539;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    3;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || |||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||| |||||||||||| ||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACCCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 553
AB114999
LOCUS       AB114999                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-15.
ACCESSION   AB114999
VERSION     AB114999.1  GI:41016507
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-15"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.6;  DB 1;  Length 540;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||| |||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        241 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```

```
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 540


RESULT 554
AB115008
LOCUS       AB115008                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-12.
ACCESSION   AB115008
VERSION     AB115008.1  GI:41016516
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-12"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.6;  DB 1;  Length 540;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGACCAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        301 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 555
AB115009
LOCUS       AB115009                540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-15.
ACCESSION   AB115009
VERSION     AB115009.1  GI:41016517
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-15"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.6;  DB 1;  Length 540;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||   | ||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 556
AB115134
LOCUS       AB115134                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE9-4.
ACCESSION   AB115134
VERSION     AB115134.1  GI:41016569
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE9-4"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.5;  DB 1;  Length 537;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   19;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              |||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AGTTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||   | ||||||||||||| ||| ||||||||||||||||||||||
Db        358 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||||| |||||||| |||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 557
AB114926
LOCUS       AB114926                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-15.
ACCESSION   AB114926
VERSION     AB114926.1  GI:41016434
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-15"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.6%;  Score 493.3;  DB 1;  Length 539;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||| |||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAATCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||| |||||||||||||||:|||:|||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||||||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGG 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| ||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539
```

RESULT 558
AB120058
LOCUS AB120058 521 bp RNA linear VRL 01-FEB-2006
DEFINITION Human endogenous retrovirus W genomic RNA, region corresponding to gag, isolate:HWG10-3.
ACCESSION AB120058
VERSION AB120058.1 GI:45330729
KEYWORDS .
SOURCE Human endogenous retrovirus W
ORGANISM Human endogenous retrovirus W
Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 521)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University, Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .521
/organism="Human endogenous retrovirus W"
/mol_type="genomic RNA"
/isolate="HWG10-3"
/isolation_source="spleen of human"
/db_xref="taxon:87786"
misc_feature 1. .521
/note="corresponds to gag"

Query Match 6.6%; Score 492.2; DB 1; Length 521;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 509; Conservative 9; Mismatches 2; Indels 2; Gaps 2;

```
Qy        632 AAACGCCCCTAAGACGTATTCTGGARAATTGGGAMCAATTTGACCCTCAGACA-CTAAGA 690
              |||||||| |||||||||||||||:||||||||:|||||||||||||||||| ||||||
Db          1 AAACGCCCGTAAGACGTATTCTGGAGAATTGGGACCAATTTGACCCTCAGACAGCTAAGA 60

Qy        691 AAGAAACGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTA 750
              |||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAAGCGACTTATATTCTTCTGCAGTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTA 120

Qy        751 TAACACCATCTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATA 810
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        121 TAACACCATCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATA 180

Qy        811 AGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATG 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTACAAACTTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATG 240

Qy        871 CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAM 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        241 CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAA 300

Qy        931 YTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACA 990
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTAATAAGGACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACA 360

Qy        991 GTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGA 1050
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        361 GTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGA 419

Qy       1051 GAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAA 1110
              |||||||||||||||||||||||||||||:||||::|||||||||||||||||||||||||
Db        420 GAATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAA 479

Qy       1111 ACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||||||||||||||:|||||||||||||||
Db        480 ACAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATT 521
```

RESULT 559
AB114939
LOCUS AB114939 539 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE2-15.
ACCESSION AB114939
VERSION AB114939.1 GI:41016447
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 539)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE2-15"
/db_xref="taxon:9606"
/tissue_type="prostate"
repeat_region 1. .539
/note="HERV-W env"

Query Match 6.6%; Score 492.1; DB 1; Length 539;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 514; Conservative 3; Mismatches 22; Indels 1; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60
```

```
Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || |||||||||| | |||||||||||| ||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACGACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| |||||||| ||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCTCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 560
AB114944
LOCUS       AB114944                 536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-13.
ACCESSION   AB114944
VERSION     AB114944.1  GI:41016452
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-13"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.6%;  Score 492;  DB 1;  Length 536;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    4;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||    || |||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAAT----ACTACTTCCTTTTTAGTAGGACCTCTTGTTTCCAAT 56

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||| |||||||||||||||||||||||||||||||||||||||||| | |||||
Db         57 CTGGAAACAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 116

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        117 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 176

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||| |||||||||||||||| |||||||||||||||| |||||
Db        177 GGAATATTTTTTGTCTGTAGTACCTCAGCCTATCGTCGTTTGAATGGCTCTTCAAAATCT 236

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:|||||||||||| ||||||||||||
Db        237 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGCACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || | ||||||| ||||||||||||||||||||||||||||||  ||
Db        297 AATCATGTCGTACCAAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 536


RESULT 561
AB114956
LOCUS       AB114956                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-8.
ACCESSION   AB114956
VERSION     AB114956.1  GI:41016464
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-8"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.6%;  Score 492;  DB 1;  Length 540;
```

```
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  507;  Conservative    4;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| |||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        241 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | |||||||||||| ||||||||||||||||| ||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        421 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||| ||||||||||||||||||:||||| ||||| ||||||||
Db        481 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 540


RESULT 562
AB114972
LOCUS       AB114972                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-2.
ACCESSION   AB114972
VERSION     AB114972.1  GI:41016480
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-2"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.6%;  Score 491.7;  DB 1;  Length 539;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    4;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| |||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        300 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||| ||| ||||||||||||||||||||||||
Db        360 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||| |||||||||||||| ||||:|||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 539


RESULT 563
AB126149
LOCUS       AB126149                539 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE6-6.
ACCESSION   AB126149
VERSION     AB126149.1  GI:47169749
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE6-6"
                     /db_xref="taxon:9542"
```

```
                     /tissue_type="pancreas"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.6%;  Score 491.7;  DB 1;  Length 539;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  513;  Conservative    4;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| ||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCACCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||||||  |||||||||||||| ||||||||||||||||||||||||  ||
Db        301 AATTATGTCATACATAAGCCCCGCAACAAGAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||||| |||||||||||| ||| ||||||||||||
Db        421 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 539


RESULT 564
AB115173
LOCUS       AB115173                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE17-5.
ACCESSION   AB115173
VERSION     AB115173.1  GI:41016608
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE17-5"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.5%;  Score 490.4;  DB 1;  Length 540;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 6318
             :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||  ||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||| || |||||||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        361 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 565
AB115170
LOCUS       AB115170                 537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE17-2.
ACCESSION   AB115170
VERSION     AB115170.1  GI:41016605
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
```

```
           of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
           Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE17-2"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 490.3;  DB 1;  Length 537;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    4;  Mismatches   21;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||  || ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||| |||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 566
AB115003
LOCUS       AB115003                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-4.
ACCESSION   AB115003
VERSION     AB115003.1  GI:41016511
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-4"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 490.1;  DB 1;  Length 539;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    4;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAGACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||| |||||||||||||||:|||:|||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| |||||||| |||||||||||||||||||||||  ||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAGGAGTACCCATTCTTCCTTTTGTTATCCGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||  |||||||    | |||||||||||||  ||| ||||||||||||||||||||
Db        360 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539


RESULT 567
AB114942
LOCUS       AB114942                541 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-6.
ACCESSION   AB114942
VERSION     AB114942.1  GI:41016450
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
```

```
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-6"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .541
                     /note="HERV-W env"

  Query Match             6.5%;  Score 489.9;  DB 1;  Length 541;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  513;  Conservative    3;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTA-CCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
              |||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATC 240

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
              | |||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||
Db        241 TGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATA 300

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
              || | |||||||| ||||||||||||||||||||||||||||||||||||||||||| ||
Db        301 CAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGG 360

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 420

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
              |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  |||||||
Db        421 CTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        481 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 540

Qy       6738 A 6738
              |
Db        541 A 541


RESULT 568
AB115111
LOCUS       AB115111                 537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE4-5.
ACCESSION   AB115111
VERSION     AB115111.1  GI:41016546
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE4-5"
                     /db_xref="taxon:9606"
                     /cell_line="HCT-116"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 489.1;  DB 1;  Length 537;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    3;  Mismatches   22;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCTCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGGATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 569
AB096023
LOCUS       AB096023                 550 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-3.
ACCESSION   AB096023
```

```
VERSION     AB096023.1  GI:30698499
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hong,K., Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of HERV-W env Fragments in Fetal Brain cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 550)
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-3"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .550
                     /gene="env"
     repeat_region   1. .550
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             6.5%;  Score 489;  DB 1;  Length 550;
  Best Local Similarity   97.8%;  Pred. No. 0;
  Matches  531;  Conservative    4;  Mismatches    4;  Indels    4;  Gaps    4;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAAGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||| |||| |||||||||||||||||||| ||||||||||||||||||||||
Db        361 GCAGGAGTGCTGGGTGTACTAGGTACTGGCATTGGCGGCATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAC-GGGTCGCCGACTCCCTGGT 6677
             ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGGTCGCCGACTCCCTGGT 480

Qy       6678 CACC-TTGCAAGATCAACTTAACTCCC-TAGCAGCAGTAGTCC-TTCRAAATCGAAGAGC 6734
             |||| |||||||||||||||||||||| ||||||||||||||| |||:||||||||||||
Db        481 CACCTTTGCAAGATCAACTTAACTCCCTTAGCAGCAGTAGTCCTTTCAAAATCGAAGAGC 540

Qy       6735 TTT 6737
             |||
Db        541 TTT 543


RESULT 570
AB114924
LOCUS       AB114924                 535 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-13.
ACCESSION   AB114924
VERSION     AB114924.1  GI:41016432
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-13"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.5%;  Score 488.9;  DB 1;  Length 535;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             | | ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db          11 ACACATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             |||||||||||||||||||||||||||||||||||| | | ||||||||||||||||||||
Db          71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
             |||||||||||||||||| ||| ||| ||||||||||||||||||||||||||||||||||
Db         131 CATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 189

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||
Db         190 CTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTGTC 249

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
             ||||||||||||||:|||:||||||||||||||||||||| ||||| | ||||||| |
Db         250 ATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATACAATCATGTCATACC 309

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 TAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 369

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         370 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 429

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         430 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 489
```

```
Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        490 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 535
```

RESULT 571
AB114935
LOCUS AB114935 539 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE2-11.
ACCESSION AB114935
VERSION AB114935.1 GI:41016443
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 539)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE2-11"
/db_xref="taxon:9606"
/tissue_type="prostate"
repeat_region 1. .539
/note="HERV-W env"

Query Match 6.5%; Score 488.9; DB 1; Length 539;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 512; Conservative 3; Mismatches 24; Indels 1; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        240 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        300 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        360 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        420 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539
```

RESULT 572
AB114951
LOCUS AB114951 539 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE5-2.
ACCESSION AB114951
VERSION AB114951.1 GI:41016459
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 539)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE5-2"
/db_xref="taxon:9606"
/tissue_type="kidney"
repeat_region 1. .539
/note="HERV-W env"

Query Match 6.5%; Score 488.9; DB 1; Length 539;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 512; Conservative 3; Mismatches 24; Indels 1; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
```

```
                 |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                 |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                 |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 573
AB114977
LOCUS       AB114977                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-11.
ACCESSION   AB114977
VERSION     AB114977.1  GI:41016485
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-11"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 488.9;  DB 1;  Length 539;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    3;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                 ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                 :|||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                 |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                 |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                 | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||| |||
```

```
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                 |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                 |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                 |||||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 574
AB115282
LOCUS       AB115282                660 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP8-5.
ACCESSION   AB115282
VERSION     AB115282.1  GI:41052510
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 660)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP8-5"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W pol"
     repeat_region   1. .660
                     /note="HERV-W pol"

  Query Match             6.5%;  Score 488.8;  DB 1;  Length 660;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  553;  Conservative   59;  Mismatches   41;  Indels   10;  Gaps    7;

Qy       3091 CCAAACSARARGCTCARCTCTG--CTCACAGCAGGTTAAATACTTAGGRCTAARA--TTA 3146
                 ||||||:|:|:| |||:|| ||   |      ||||| ||||||||||:||||:|  |
Db          2 CCAAACCAAAAGGTCAGCTTTGGCTTTCAGCAAGGTTTAATACTTAGGGCTAAAAATTAT 61

Qy       3147 TCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCAT 3206
                 |||||||||||:|||||||||||||:::|||||||||||||||||||||||||||||||
Db         62 CCCAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCAT 121

Qy       3207 -CYCAAAACCCTAAAGCAACTAAG-RGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWAT 3264
                 |:||||||  |||||||||||||| :|::|||||||||:|||:|||:||||| ||||:||
Db        122 CCCCAAAACTTTAAAGCAACTAAGAGGGGTTCCTTGGCATAACAGGTTTCTGTCGAATAT 181

Qy       3265 GGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAG 3324
                 ||||| |||||||: || :|||||| ||| :|||| :|||||:|||||||||||||  |
Db        182 GGATT-CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGG 240

Qy       3325 AAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCT 3384
                 |||| ||||||||||||:||||||||:|:||||||:::||| ||||||||||||||
```

```
Db        241 AAAGTCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTA 300

Qy       3385 AAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTT 3444
              ||  | | |||  ||||||||:||||||||||||:||||| | :|||||||||||||:||
Db        301 AAGAAGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTT 358

Qy       3445 YATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYG 3504
              :|||::||| | |||||| |||:||:||||||:||||||||||||||||:|||:||||||:|
Db        359 TATATGTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACG 417

Qy       3505 AGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYT 3564
              ||||||||||::||||::|||||| ||||||||:|||||||||||||||||||||:|:|
Db        418 AGCTTGCAACCCGTGGCATACCTGAATAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCT 477

Qy       3565 CATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATA 3624
              |||||||||:|||||||||  ||||||||||::|||||||||||||||| ||||||||
Db        478 CATTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTTAAATAATA 537

Qy       3625 CAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAA 3684
              |||||:|||||| |||||||||||||||||||||:|||||::|||||||| |||||||||
Db        538 CAGGGAAGAGATTTTACTGTGTGGACATCTCATGATGTGAACGGCATACTGACTGCTAAA 597

Qy       3685 GGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGG 3744
              |||||||||||||||||||||||:||||||||||||:||||||||||||||||||:||
Db        598 GGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGG 657

Qy       3745 CCA 3747
              |||
Db        658 CCA 660


RESULT 575
AB126160
LOCUS       AB126160                539 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-7.
ACCESSION   AB126160
VERSION     AB126160.1  GI:47169760
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-7"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 488.5;  DB 1;  Length 539;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    4;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||| ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||| ||:|||:|||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCACCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||||| |||||||||||||| ||||||||||||||||||||||||||  ||
Db        301 AATTATGTCATACATAAGCCCCGCAACAAGAGAGTACCCATTCTTCCTTTTGTTATTAGA 360

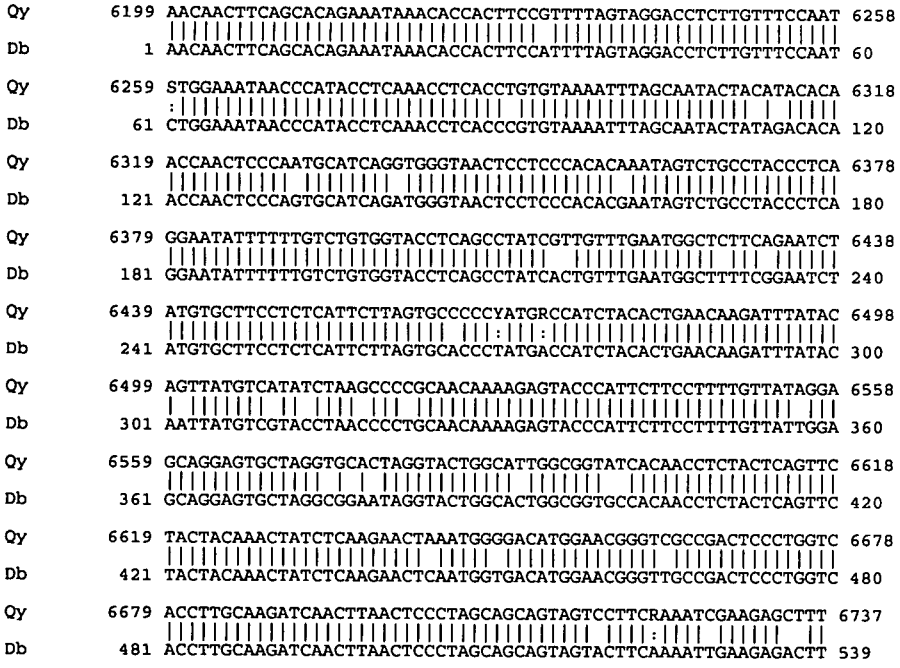
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||| |||||||||||||||||| |||||||||||| ||| ||||||||||
Db        421 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| |||||||||| ||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 539


RESULT 576
AB115128
LOCUS       AB115128                541 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE8-3.
ACCESSION   AB115128
VERSION     AB115128.1  GI:41016563
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE8-3"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W env"
     repeat_region   1. .541
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.8;  DB 1;  Length 541;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  504;  Conservative    4;  Mismatches   31;  Indels    0;  Gaps    0;
```



```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || |||| ||| |||||||||||||||||||||||||||||||||||| |||
Db        301 AATTATGTCGTACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| ||||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTGCCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| |||||||||||||| ||||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             |||||||||||||||||||||||||||||||||||||||||||| ||||:|||| ||||||  ||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATTGAAGAGACTT 539
```

```
RESULT 577
AB115118
LOCUS       AB115118                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE6-3.
ACCESSION   AB115118
VERSION     AB115118.1  GI:41016553
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE6-3"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"
```

```
  Query Match             6.5%;  Score 487.5;  DB 1;  Length 537;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   23;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||  ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||| |||||||| |||||| || |||||||||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 578
AB115172
LOCUS       AB115172                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE17-4.
ACCESSION   AB115172
VERSION     AB115172.1  GI:41016607
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE17-4"
                     /db_xref="taxon:9606"
```

```
                     /cell_line="U-937"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.5;  DB 1;  Length 537;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   23;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
               |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 579
AB114917
LOCUS       AB114917               539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-2.
ACCESSION   AB114917
VERSION     AB114917.1  GI:41016425
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-2"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.3;  DB 1;  Length 539;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCCGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||| |||||||||||||||| ||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             || |||||||||||||||||||||||||||||||||||||:|||||  | |||||||
Db        480 ACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTTTA 539


RESULT 580
AB114949
LOCUS       AB114949               539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE4-6.
ACCESSION   AB114949
VERSION     AB114949.1  GI:41016457
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
```

```
          Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE4-6"
                     /db_xref="taxon:9606"
                     /tissue_type="heart"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.3;  DB 1;  Length 539;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTGGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||:|||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || |||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 581
AB115151
LOCUS       AB115151                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE13-3.
ACCESSION   AB115151
VERSION     AB115151.1  GI:41016586
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE13-3"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.3;  DB 1;  Length 539;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || |||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCCTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 582
AB051009
LOCUS       AB051009                540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-20-9.
ACCESSION   AB051009
VERSION     AB051009.1  GI:14020837
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
```

```
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-20-9"
                     /db_xref="taxon:9606"
                     /chromosome="20"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47561.1"
                     /db_xref="GI:14020838"
                     /translation="NNFSTEINTTSILIGPLVSNLEITHTSNLTCVKFSNTIDTPNSQ
                     CIRWVSHPTRIVCLPSGIFFVCGTSAYRCFKGSSESICFLSFLVPPMTIYTEQDLYNH
                     VVPKPRNKRVPILTFVVGAGGLGGLGTGIGGITTSTQFYYKLSQELNGDMEWVADSLV
                     TLQDQLNSLVAVVLQNRRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             6.5%;  Score 487.2;  DB 1;  Length 540;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  504;  Conservative    4;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAATAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||| || |||||||| ||||||||||||||||||||
Db        121 CCCAACTCCCAATGCATCAGGTGGGTAAGTCATCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||| || ||||| |||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTCAAGGGCTCATCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              || |||||||||||||||||||||||:|||:|||||||||||||||| |||||||||
Db        241 ATATGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAGGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || |||||||||||||||||||||||||||||||| |||||||| | |||
Db        301 AATCATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTACTTTGTTGTGGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||| | |||| | | |||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGGGTTAGGCGGATTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||| ||||||||| |||| || |||||||| |||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAATGGGTTGCTGACTCCCTAGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||| ||||||||| |||||:|||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGTAGCAGTAGTTCTTCAAAATCGAAGAGCTTTA 540

RESULT 583
AB114920
LOCUS       AB114920                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-7.
ACCESSION   AB114920
VERSION     AB114920.1  GI:41016428
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-7"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.5%;  Score 487.2;  DB 1;  Length 540;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  504;  Conservative    4;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db          1 AACAACTTCAGCACAGAAATAATCACCACTTCCGTTTTAGTAGGACCTCTTGCTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||| |||||||||||||||||||||||||||||||||||| |||| | |||||
Db         61 GTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAACACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||| |||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCTTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||| |||||||||| |||||||||| ||| ||||| ||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCATTGTTTGAATAGCTTTTCAGTATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||||||  ||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAATGAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||  || |||||||||||||||||||||| |||||||||||||||||||| |||
Db        301 AATCATGTTGTACCTAAGCCCCGCAACAAAAGAGTTCCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||  ||| || ||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGTTAGCTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||| |||||||||||||||||||| |||||||| |||| ||| ||||||||||||
Db        421 TACTACCAACTATCTCAAGAACTAAATGGTGACATGGAATGGGTTGCCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||| |||||:|||||||||| ||||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGGCCTTCAAAATCGAAGAGATTTA 540
```

```
RESULT 584
AF127227
LOCUS       AF127227                544 bp    mRNA    linear   VRL 11-AUG-1999
DEFINITION  Multiple sclerosis associated retrovirus element clone CL4 env-like
            mRNA sequence.
ACCESSION   AF127227
VERSION     AF127227.1  GI:5726292
KEYWORDS    .
SOURCE      Multiple sclerosis associated retrovirus element
  ORGANISM  Multiple sclerosis associated retrovirus element
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F.,
            Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E.,
            Mallet,F., Mandrand,B. and Perron,H.
  TITLE     Molecular cloning and characterization of MSRV-related sequences
            associated with retrovirus-like particles
  JOURNAL   Virology 260 (1), 1-9 (1999)
   PUBMED   10405350
REFERENCE   2  (bases 1 to 544)
  AUTHORS   Komurian-Pradel,F., Paranhos-Baccala,G., Bedin,F.,
            Ounanian-Paraz,A., Sodoyer,M., Ott,C., Rajoharison,A., Garcia,E.,
            Mallet,F., Mandrand,B. and Perron,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-FEB-1999) UMR103 CNRS, bioMerieux, 46, allee
            D'Italie, Lyon 69007, France
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Multiple sclerosis associated retrovirus
                     element"
                     /mol_type="mRNA"
                     /db_xref="taxon:89382"
                     /clone="CL4"
                     /note="isolated in cell-free plasma from patient with MS"
     misc_feature    1. .544
                     /note="similar to env but interrupted by premature stop
                     codon and frameshift"

  Query Match             6.5%;  Score 487.1;  DB 1;  Length 544;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  512;  Conservative    4;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       6069 CCTGGTAAGCCTATTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCC 6128
             |||||| |||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db          1 CCTGGTGAGCCTATTTAATACCACCCTCACTGGGCTCCATGACGTCTCAGCCCAAAACCC 60

Qy       6129 TACTAACTGTTGGATATGCCTCCCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACC 6188
             |||||||||||||| |||||||||| ||||||:| |||| ||||||||||||||||
Db         61 TACTAACTGTTGGATGTGCCTCCCCCTGCACTTCAGGTCATACATTTCAATCCCTGTACC 120

Qy       6189 TGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCT 6248
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        121 TGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTGGTAGGACCTCT 180

Qy       6249 TGTTTCCAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATAC 6308
             ||||||||||:|||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 TGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTATGTAAAATTTAGCAATAC 240

Qy       6309 TACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTG 6368
             || | |||||||||||||||||||||||||||||  |||||||||||||  |||||||||
Db        241 TATAGACACAACCAACTCCCAATGCATCAGGTGAATAACTCCTCCCACAGGAATAGTCTG 300

Qy       6369 CCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTC 6428
             ||||||| |||||||| |||||||||||||| |||||||||||||||||||||||||||||
Db        301 CCTACCCCCAGGAATA-TTTTTGTCTGTGGCACCTCAGCCTATCGTTGTTTGAATGGCTC 359

Qy       6429 TTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACA 6488
             ||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||
Db        360 TTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACA 419

Qy       6489 AGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTT 6548
             |||||||||| ||||||  || |||||||| ||||||||||||||||||||||||||||||
Db        420 AGATTTATACAATTATGTTGTACCTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTTT 479

Qy       6549 TGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTC 6608
             |||||| |||||||||||||||||||| ||||||| |||||||||| ||||||||||||
Db        480 TGTTATCAGAGCAGGAGTGCTAGGTGGACTAGGTGCTGGCATTGGCAGTATCACAACCTC 539

Qy       6609 TACTC 6613
             |||||
Db        540 TACTC 544


RESULT 585
AB126142
LOCUS       AB126142                539 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE3-3.
ACCESSION   AB126142
VERSION     AB126142.1  GI:47169742
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:825158129620)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE3-3"
                     /db_xref="taxon:9542"
                     /tissue_type="kidney"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 486.9;  DB 1;  Length 539;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||| ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||| ||:|||:||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCACCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||||||  |||||||||||||||| ||||||||||||||||||||||||  ||
```

```
Db        301 AATTATGTCATACATAAGCCCCGCAACAAGAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||| ||  |||||||||| ||| ||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGATGTGGAACGGGTTGCCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 539


RESULT 586
AB126137
LOCUS       AB126137                537 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE2-4.
ACCESSION   AB126137
VERSION     AB126137.1  GI:47169737
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE2-4"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 486.1;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||| |||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| |||||||: ||:|||||||||||||||||||||||| ||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGATTTAAAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||| || ||||||||| |||||||||||||||||||||||||||||||| |||
Db        298 AGTTATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | |||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||  |||||||| |||| | ||||||||||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTAACATGGAACAGGTCACTGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 587
AB051005
LOCUS       AB051005                537 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-12-2.
ACCESSION   AB051005
VERSION     AB051005.1  GI:14020831
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
  PUBMED    11375061
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-12-2"
                     /db_xref="taxon:9606"
                     /chromosome="12"
     repeat_region   1. .537
                     /rpt_family="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| |||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||| ||||||||| ||||||| || ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 588
AB051008
LOCUS       AB051008                537 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-17-5.
ACCESSION   AB051008
VERSION     AB051008.1  GI:14020835
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-17-5"
                     /db_xref="taxon:9606"
                     /chromosome="17"
     gene            1. .537
                     /gene="env"
     CDS             <1. .>537
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47560.1"
                     /db_xref="GI:14020836"
                     /translation="NNFSTEINTTSVLVGPLSNLEITHTSNLTCVKFSNTIDTANSQC
                     IRWVTPPTRIVCLPSGIFFVCGTSAYHCLNGSSESVCFLSFLVAPMPIYTEQDLYNHV
                     IPKPRNKRVPILPFVIGAGVLGGVATGIGGITTSTQFYYKLSQELNGDMEWVADTLVT
                     LQDQLNSLAAVVLQNRRAL"
     repeat_region   1. .537
                     /rpt_family="HERV-W"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| |||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| ||||||||||||||| |||||||| ||||||| || ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 589
AB114927
LOCUS       AB114927                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE1-16.
ACCESSION   AB114927
VERSION     AB114927.1  GI:41016435
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE1-16"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
```

```
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||   ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 590
AB114932
LOCUS       AB114932                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-1.
ACCESSION   AB114932
VERSION     AB114932.1  GI:41016440
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-1"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||   ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 591
AB114938
LOCUS       AB114938                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-14.
ACCESSION   AB114938
VERSION     AB114938.1  GI:41016446
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-14"
```

```
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGGATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                 |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 592
AB114941
LOCUS       AB114941                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE3-5.
ACCESSION   AB114941
VERSION     AB114941.1  GI:41016449
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE3-5"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||| ||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACTTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                 |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 593
AB114974
LOCUS       AB114974                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-4.
ACCESSION   AB114974
VERSION     AB114974.1  GI:41016482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
```

```
           Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-4"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 594
AB114983
LOCUS       AB114983                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-7.
ACCESSION   AB114983
VERSION     AB114983.1  GI:41016491
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-7"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 595
AB114994
LOCUS       AB114994                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-5.
ACCESSION   AB114994
VERSION     AB114994.1  GI:41016502
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
```

```
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-5"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||| ||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 596
AB115098
LOCUS       AB115098                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE1-5.
ACCESSION   AB115098
VERSION     AB115098.1  GI:41016533
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE1-5"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| |||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||| ||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 597
AB115101
LOCUS       AB115101                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE2-3.
ACCESSION   AB115101
VERSION     AB115101.1  GI:41016536
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE2-3"
                     /db_xref="taxon:9606"
                     /cell_line="PFSK-1"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        418 TACTACAAACTGTCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 598
AB115106
LOCUS       AB115106                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE3-3.
ACCESSION   AB115106
VERSION     AB115106.1  GI:41016541
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE3-3"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 599
AB115109
LOCUS       AB115109                537 bp    RNA     linear   PRI 01-FEB-2006
```

```
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE4-1.
ACCESSION   AB115109
VERSION     AB115109.1  GI:41016544
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE4-1"
                     /db_xref="taxon:9606"
                     /cell_line="HCT-116"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 600
AB115127
LOCUS       AB115127                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE8-2.
ACCESSION   AB115127
VERSION     AB115127.1  GI:41016562
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE8-2"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477
```

```
Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 601
AB115131
LOCUS       AB115131                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE9-1.
ACCESSION   AB115131
VERSION     AB115131.1  GI:41016566
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE9-1"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCC 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||  ||||||||||||||||| |||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGCCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 602
AB115149
LOCUS       AB115149                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE12-5.
ACCESSION   AB115149
VERSION     AB115149.1  GI:41016584
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE12-5"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357
```

```
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 603
AB115155
LOCUS       AB115155                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE14-1.
ACCESSION   AB115155
VERSION     AB115155.1  GI:41016590
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE14-1"
                     /db_xref="taxon:9606"
                     /cell_line="LOX-IMVI"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 604
AB115163
LOCUS       AB115163                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE15-4.
ACCESSION   AB115163
VERSION     AB115163.1  GI:41016598
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE15-4"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.9;  DB 1;  Length 537;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 605
AB115095
LOCUS       AB115095                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE1-2.
ACCESSION   AB115095
VERSION     AB115095.1  GI:41016530
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE1-2"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.7;  DB 1;  Length 539;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 606
AB115148
LOCUS       AB115148                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE12-4.
ACCESSION   AB115148
VERSION     AB115148.1  GI:41016583
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE12-4"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.7;  DB 1;  Length 539;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    3;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAGTAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||| ||||:|||| ||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||  || |||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        300 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||| |||| ||||||| |||||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTGGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 607
AB051001
LOCUS       AB051001                 540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-5-6.
ACCESSION   AB051001
VERSION     AB051001.1  GI:14020826
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-5-6"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     repeat_region   1. .540
                     /rpt_family="HERV-W env"

  Query Match             6.5%;  Score 485.6;  DB 1;  Length 540;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||| ||||||| |||||||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCCCAGAAATGAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||| |||||||||||||||||||||||||||||||||| | ||| |
Db         61 CTGGAAATAACCCATATCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAAA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              || |||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        121 ACAAACTCCCAAGGCATCAGGTGGGTAACTGCTCCCACACAAATAGTCTGCCTACCTTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| |||||||||||  |||||||||  |
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGACTCTTCAGAAAGT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||||:| |:||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTACGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||||||  |||||||||||||||||||||||||||||||||||||| |||||| |||
Db        301 AATTATGTCATACCTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTCTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTAGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||  |||||||||||||||| |||||||| | |||| || ||||||||||
Db        421 TACTACAAACTATTTCAAGAACTAAATGGTGACATGGAGCAGGTCACCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||| | |||||||||||||||||||||||||||||||||||| :|||| ||||||||||
Db        481 ACCTAGGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTTAAAATTGAAGAGCTTTA 540


RESULT 608
AB114984
LOCUS       AB114984                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-8.
ACCESSION   AB114984
VERSION     AB114984.1  GI:41016492
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-8"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.6;  DB 1;  Length 540;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
```

```
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| ||||||| | ||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGTAATACTATAGACACG 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||| |||||||||||||||||||||||||||||||||||||||| |||||||||||||| |
Db        121 ACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTTA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||| ||||| || |||||||||||||| ||||||||| ||||||| ||
Db        181 GGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAGAACCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||:|||:|||| |||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||||| ||||||||| |||||||| ||||| |||||||||||||||||  ||
Db        301 AATTATGTCATACCTAAGCCCCACAACAAAACAGTACTCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  ||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||| || |||||||||| |||| ||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||| |||||||||| ||||||||||:|||| ||||||||||
Db        481 ACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 609
AB114992
LOCUS       AB114992                 540 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-3.
ACCESSION   AB114992
VERSION     AB114992.1  GI:41016500
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-3"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.6;  DB 1;  Length 540;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||| ||||||||||||||||:|||:|||||||||||||||||||| ||||
Db        241 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | |||||||||||| |||||||||||||||| ||||||||||
Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| ||||||| ||||||| |||||||||| |||||||| ||||||||
Db        421 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||  |||||||||||||||| ||||||||||||||||||||:||||| ||||| ||||
Db        481 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 540


RESULT 610
AB114933
LOCUS       AB114933                 537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-3.
ACCESSION   AB114933
VERSION     AB114933.1  GI:41016441
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-3"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .537
                     /note="HERV-W env"
```

```
  Query Match             6.5%;  Score 485.5;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||  |||||||||||||| ||| ||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||:|||:||||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||    | ||||||||||||| ||| ||||||||||||||||||||||
Db        358 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:|||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 537


RESULT 611
AB115147
LOCUS       AB115147                 537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE12-3.
ACCESSION   AB115147
VERSION     AB115147.1  GI:41016582
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE12-3"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.5;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   24;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        118 ACCAGCTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || ||||||||| |||||||||||||||||||||||||||||||  ||
Db        298 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||    |||||||||||||||||| ||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||||||
Db        418 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 612
AB114980
LOCUS       AB114980                 539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-14.
ACCESSION   AB114980
VERSION     AB114980.1  GI:41016488
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
```

```
                     /mol_type="other RNA"
                     /isolate="WE9-14"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.3;  DB 1;  Length 539;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| ||||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| |||||||||||||||| ||||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| ||||||| ||||||| |||||||||| ||||||||| ||||||||||
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||| |||||||||||||||||||:||||| |||||| ||||
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 613
AB115158
LOCUS       AB115158                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE14-4.
ACCESSION   AB115158
VERSION     AB115158.1  GI:41016593
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)

FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE14-4"
                     /db_xref="taxon:9606"
                     /cell_line="LOX-IMVI"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 485.3;  DB 1;  Length 539;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||| ||||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACT-CTCCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| ||||||||||||||||||:|||:|||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||| |||||||| |||||||| ||||||| ||  ||||||||
Db        420 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539


RESULT 614
AB115135
LOCUS       AB115135                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE9-5.
ACCESSION   AB115135
VERSION     AB115135.1  GI:41016570
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE9-5"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.4;  DB 1;  Length 536;
  Best Local Similarity   94.5%;  Pred. No. 0;
  Matches  497;  Conservative    4;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             | | ||||||||||||||| |||||||||||||||||||||||||||:|||||||||||||
Db         11 ACACATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
             ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        131 CATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 190

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             |||||||||||||||||||  |||||||||||||||||||||||| ||||||||||||||
Db        191 CTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTC 250

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
             |||||||||||  |||:|||:|||||||||||||||||||||||||||||  ||||||  || |
Db        251 ATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTTGTACC 310

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             |||||||| |||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db        311 TAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGG 370

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
             || ||||||| |||||||||||||| |||||||||||||||||||||||||||||||||||
Db        371 TGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACTACAAACTATC 430

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             ||||||||| ||||| ||||||||| |||| ||||||||||||||||||||||||||||||
Db        431 TCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCA 490

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||| ||||||| ||||||||||:|||| ||||||||||||
Db        491 ACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 615
AB055413
LOCUS       AB055413                 537 bp    DNA     linear   PRI 07-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:AZ521-3.
ACCESSION   AB055413
VERSION     AB055413.1  GI:15128495
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Lee,W.-H., Kim,H.-M. and Kim,H.-S.
  TITLE     Identification of new endogenous retroviral sequences belonging to
            the HERV-W family in human cancer cells
  JOURNAL   Intervirology (2001) In press
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="AZ521-3"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   |||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGCGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCG 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 616
AB055415
LOCUS       AB055415                 537 bp    DNA     linear   PRI 07-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:OVCAR-3.
ACCESSION   AB055415
VERSION     AB055415.1  GI:15128497
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.-M., Lee,W.-H., Kim,H.-M. and Kim,H.-S.
  TITLE     Identification of new endogenous retroviral sequences belonging to
            the HERV-W family in human cancer cells
  JOURNAL   Intervirology (2001) In press
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="OVCAR-3"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGCAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GCGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 617
AB114950
LOCUS       AB114950                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-1.
ACCESSION   AB114950
VERSION     AB114950.1  GI:41016458
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-1"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        298 TATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | ||| |||||||||||||||||||||| ||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAGCCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 618
AB115004
LOCUS       AB115004                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-6.
ACCESSION   AB115004
VERSION     AB115004.1  GI:41016512
```

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-6"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| |||||||| |||| ||||||||||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 619
AB115105
LOCUS       AB115105                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE3-2.
ACCESSION   AB115105
VERSION     AB115105.1  GI:41016540
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE3-2"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| |||||||| ||||| || ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||| ||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        478 ACCTTACAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 620
AB115161
LOCUS       AB115161                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE15-2.
ACCESSION   AB115161
VERSION     AB115161.1  GI:41016596
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE15-2"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGCGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||| |||||||||||:|||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCGGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 621
AB115165
LOCUS       AB115165                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE16-1.
ACCESSION   AB115165
VERSION     AB115165.1  GI:41016600
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE16-1"
                     /db_xref="taxon:9606"
                     /cell_line="2F7"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 622
AB115175
LOCUS       AB115175                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE18-2.
ACCESSION   AB115175
VERSION     AB115175.1  GI:41016610
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE18-2"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| |||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGCCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| |||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 623
AB115176
LOCUS       AB115176                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE18-3.
ACCESSION   AB115176
VERSION     AB115176.1  GI:41016611
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE18-3"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 484.3;  DB 1;  Length 537;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297
```

```
Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        358 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        418 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 537


RESULT 624
AB115145
LOCUS       AB115145                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE12-1.
ACCESSION   AB115145
VERSION     AB115145.1  GI:41016580
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE12-1"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.5%;  Score 483.9;  DB 1;  Length 537;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   25;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        118 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTTTTCGGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537


RESULT 625
AB115010
LOCUS       AB115010                538 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-16.
ACCESSION   AB115010
VERSION     AB115010.1  GI:41016518
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 538)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-16"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .538
                     /note="HERV-W env"

  Query Match             6.5%;  Score 483.8;  DB 1;  Length 538;
  Best Local Similarity   95.4%;  Pred. No. 0;
  Matches  515;  Conservative    3;  Mismatches   20;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
```

```
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 298

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        299 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 358

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        359 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 418

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        419 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 478

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        479 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 538


RESULT 626
AB054086
LOCUS       AB054086                539 bp    DNA     linear   PRI 06-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:HepG2-2.
ACCESSION   AB054086
VERSION     AB054086.1  GI:15005016
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
  TITLE     Molecular cloning and phylogenetic analysis of new human endogenous
            retrovirus HERV-W family in cancer cells
  JOURNAL   Curr. Microbiol. (2001) In press
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Lee,J. and Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HepG2-2"
                     /db_xref="taxon:9606"
                     /cell_line="HepG2"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.5%;  Score 483.7;  DB 1;  Length 539;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   27;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        241 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        301 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        361 GCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        421 TACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        481 ACCTTGCAAGATCATCTT-ATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 627
AB115005
LOCUS       AB115005                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-7.
ACCESSION   AB115005
VERSION     AB115005.1  GI:41016513
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-7"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.5%;  Score 483.7;  DB 1;  Length 539;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   27;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
```

```
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:|||:|||||||||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTG-CCCCTATGACCATCTACACTGAACAAAATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||   | |||||||||| |||||||||||||| |||||||||||||||||| |||
Db        300 AATCATGTTGCACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 628
AB115293
LOCUS       AB115293                652 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP11-3.
ACCESSION   AB115293
VERSION     AB115293.1  GI:41052521
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various humán tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 652)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP11-3"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W pol"
     repeat_region   1. .652
                     /note="HERV-W pol"

  Query Match             6.5%;  Score 483.3;  DB 1;  Length 652;
  Best Local Similarity   82.3%;  Pred. No. 0;
  Matches  544;  Conservative   57;  Mismatches   49;  Indels   11;  Gaps    6;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
              |||||||:|:|:|||||:|||||||||| |||||||||||||| ||||:  ||:|||||||


Db          1 TCCAAACCAAAGGCTCAGCTCTGCTCATAGCAGGTTAAATACGTAGGGGCAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
              ||||||||||:||| ||||||||||||:::|||||||||||||||||||| ||||||||:|
Db          61 AAAGGCACCAGGGCACTCAGTGAGGAACGTATCCAGCCTATACTGGCTTGTCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
              |||||||||||||||||| :|:: |||||||:|||:|||:|||||||||:|||||||
Db         121 AAAACCCTAAAGCAACTAAAAGGGCTCCTTGGCATAACAGGTTTCTGCCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
              |||||||: |||:|||||||||| |||||:|     |||||||||||||||||||||
Db         181 -CCCAGGTACGGCGAAATAGCCAGAACATTATA-----TAATTAAGGAAACTCAGAAAGC 234

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
              |||||||||||||:|||||||||:|:|||| |:::|||  |||||||||| |  ||  |
Db         235 CAATACCCATTTAGTAAGATGGACACCTGAGGCAGAAGCAGCTTTCCAGGCTCTAAAGAA 294

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
              | |||  |||||||:|||||  ||||||:|||||||:|| |||||||||:||:|||
Db         295 GGCCCT--AACCCAAGCCCCAGCGTTAAGCTTGCCAACGGGACAAGACTTTTCTTTATAA 352

Qy       3450 RTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
              :|||||| |||||||||:||:||||||:||||||||||||||:|||:|||||:||||||
Db         353 GTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAGCTT 411

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
              ||||||::||||::||||||:||||||||:|||||||||||||||| ||||:|: |||||
Db         412 GCAACCCATGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGAGTTGGCCACATTG 471

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTT-AAAATAATACAGG 3628
              ||||:||||| ||| | |||||||||::|||||||||||    ||| ||||| |||||||
Db         472 TTTACGGGTATTGGCGACAGTAGCAGTCTTAGTATCTGAAAAGGTTAAAAATGATACAGG 531

Qy       3629 GRAGAGAT-CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
              |:|||||| |||||||| ||||||||||||:|||||::|||||||||||||||||||||
Db         532 GAAGAGATCCTTACTGTGGGGACATCTCATGATGTGAATGGCATACTCACTGCTAAAGGA 591

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
              ||||| |||||||| |||||:| |||||| |||:|  ||||   ||   || |:| |||
Db         592 AACTTGGGGCTGTCAAACAACTGGTTACTTTAATATAGGGCTTATTTTACTGGAGGCCCA 651

Qy       3748 G 3748
              |
Db         652 G 652


RESULT 629
AB096034
LOCUS       AB096034                538 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-10.
ACCESSION   AB096034
VERSION     AB096034.1  GI:30698510
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 538)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .538
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-10"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .538
                     /gene="env"
     repeat_region   1. .538
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.4%;  Score 483;  DB 1;  Length 538;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches  514;  Conservative    4;  Mismatches   19;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | |||||||||||||  |||||||||||||| | ||||||||||||
Db        361 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
               |||||||| |||||||||||||||| |||||||| | |||| || ||||||||||||
Db        420 --CTACAAACTGTCTCAAGAACTAAATGGTGACATGGAGCAGGTCACCAACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACT-CCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             |||| | ||| |||||||||| |||||||||||||||||||| :|||| ||||||||||
Db        478 ACCTAGGAAGGTCAACTTAACTCCCCTAGCAGCAGTAGTCCTTTAAAATTGAAGAGCTTT 537

Qy       6738 A 6738
              |
Db        538 A 538


RESULT 630
AB051000
LOCUS       AB051000                 540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-4-1.
ACCESSION   AB051000
VERSION     AB051000.1  GI:14020824
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-4-1"
                     /db_xref="taxon:9606"
                     /chromosome="4"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47557.1"
                     /db_xref="GI:14020825"
                     /translation="NNFSTEINTTSVLVGPLVSNLEITHTSNLTCVKFSNTIDTTNSQ
                     CIRWVTPPTQIVCLPSGIFFVCGTSAYHCLNGSSGSMWFLSFLAPPMIIYTEQDLYNY
                     VIPKSHNKRVPILSFVIGAGVLDGLGAGIGGIKTSTQFYYKLSQELNGDMDWVTNSLV
                     TLQDQLNSLAAVVPQNRRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             6.4%;  Score 482.8;  DB 1;  Length 540;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    3;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             |||| |||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 TTGGAGATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||  |||
Db        181 GGAATATTTTTTGTTTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGGGTCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||| |||||||||||||||| ||||||:|||: ||||||||||||||||||||||||||
Db        241 ATGTGGTTCCTCTCATTCTTAGCGCCCCCTATGATCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||||| ||||| |||  |||||||||||||||||||| |||||||||| |||
Db        301 AATTATGTCATACCTAAGTCCCATAACAAAAGAGTACCCATTCTTTCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||  | |||||||| |||||||||| ||||||| ||||||||||||||||
Db        361 GCAGGAGTGCTAGACGGACTAGGTGCTGGCATTGGTGGTATCAAAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||| |||||||||||||||||||||| ||||||||  ||||| || |||| ||||||
Db        421 TACTATAAACTATCTCAAGAACTAAATGGTGACATGGATTGGGTCACCAACTCACTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| |||||||||||||||| ||:|||||||||||||||||
Db        481 ACCTTGCAAGATCAACTTAACTCTCTAGCAGCAGTAGTCCCTCAAAATCGAAGAGCTTTA 540


RESULT 631
```

```
AB115002
LOCUS       AB115002                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE12-3.
ACCESSION   AB115002
VERSION     AB115002.1  GI:41016510
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE12-3"
                     /db_xref="taxon:9606"
                     /tissue_type="uterus"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.7;  DB 1;  Length 537;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   26;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGGATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||| ||||||| ||  ||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 632
AB115100
LOCUS       AB115100                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE2-2.
ACCESSION   AB115100
VERSION     AB115100.1  GI:41016535
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE2-2"
                     /db_xref="taxon:9606"
                     /cell_line="PFSK-1"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.7;  DB 1;  Length 537;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   26;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||| ||||||| ||||||| ||||  |||||||||
Db        418 TACTACAAACTGTCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 477
```

```
Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537
```

```
RESULT 633
AB115115
LOCUS       AB115115                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE5-4.
ACCESSION   AB115115
VERSION     AB115115.1  GI:41016550
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE5-4"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.7;  DB 1;  Length 537;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   26;  Indels    3;  Gaps    1;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||| |||||||||| ||||    |||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---CTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||  | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||||| ||||||||| ||||||| ||   ||||||||
Db        418 CACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 537
```

```
RESULT 634
AB115167
LOCUS       AB115167                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE16-3.
ACCESSION   AB115167
VERSION     AB115167.1  GI:41016602
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE16-3"
                     /db_xref="taxon:9606"
                     /cell_line="2F7"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.7;  DB 1;  Length 537;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   26;  Indels    3;  Gaps    1;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| |||||||||||||||||||||| |||||||||| ||||    ||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357
```

```
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||| |||||| | | ||| ||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGCGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||| |||||||| ||||||||||||||||||||
Db        418 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 635
AB114964
LOCUS       AB114964                545 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-6.
ACCESSION   AB114964
VERSION     AB114964.1  GI:41016472
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 545)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-6"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .545
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.7;  DB 1;  Length 545;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches  534;  Conservative    4;  Mismatches    2;  Indels    5;  Gaps    5;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAA-CCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
              :|||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCCTCACCTGTGTAAAATTTAGCAATACTACATACAC 120

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACT-CCTCCCACACAAATAGTCTGCCTA-CCC 6375
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db        121 AACCAACTCCCAATGCATCAGGTGGGTAACTCCCTCCCACACAAATAGTCTGCCTACCCC 180

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTA-CCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        181 TCAGGAATATTTTTTGTCTGTGGTACCCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 240

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db        241 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTT 300

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         301 ATACAGTTATGTCATATCTAAGCCCCGCAATAAAAGAGTACCCATTCTTCCTTTTGTTAT 360

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 420

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCC-C 6673
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |
Db         421 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACAGGGAACGGGTCGCCGACTCCTC 480

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db         481 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAG 540

Qy       6734 CTTTA 6738
             |||||
Db         541 CTTTA 545


RESULT 636
AB050999
LOCUS       AB050999                540 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-3-8.
ACCESSION   AB050999
VERSION     AB050999.1  GI:14020822
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-3-8"
                     /db_xref="taxon:9606"
                     /chromosome="3"
     gene            1. .540
                     /gene="env"
     CDS             <1. .>540
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47556.1"
                     /db_xref="GI:14020823"
                     /translation="NNFSTEINTTSILVGPLVSNLEITHTSNLTCVKFSNTIDTTKSQ
                     CIRWVTPPTRIVCLPLGIFFVFGSLAYRCLNGFSEPMCFLSFLVPPMTIYTEQYLYNY
                     VIPKPHNKTVLILPFVIRAGVLGGLSTGIGGITTSTQFYYKLSQELKGDMEQVTDSLV
                     TLQDHHNSLAAVVLQNQRAL"
     repeat_region   1. .540
                     /rpt_family="HERV-W env"

  Query Match             6.4%;  Score 482.4;  DB 1;  Length 540;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   35;  Indels    0;  Gaps    0;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||| ||||||| | ||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGTAATACTATAGACACG 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||| |||||||||||||||||||||||||||||||| |||||||||||||||| |
Db        121 ACCAAATCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTTA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||| ||||| || |||||||||||||| ||||||||| ||||||| ||
Db        181 GGAATATTTTTTGTCTTTGGTAGCTTAGCCTATCGTTGTCTGAATGGCTTTTCAGAACCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||| |||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATTTACACTGAACAATATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||||| ||||||||| |||||| ||||||||||||||||||  ||
Db        301 AATTATGTCATACCTAAGCCCCACAACAAAACAGTACTCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  ||| ||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGTGGGCTAAGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||| || |||||||||| |||| ||||||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAAGGGTGACATGGAACAGGTCACCGACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||| | ||| |||||||||| ||||||||||:||||| ||||||||||
Db        481 ACCTTGCAAGATCATCATAATTCCCTAGCAGCGGTAGTCCTTCAAAATCAAAGAGCTTTA 540
```

RESULT 637
AB096038
LOCUS AB096038 540 bp DNA linear PRI 14-MAY-2003
DEFINITION Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-17.
ACCESSION AB096038
VERSION AB096038.1 GI:30698514
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shin,K., Yi,J. and Kim,H.
TITLE Isolation of Human Endogenous Retrovirus HERV-W env Fragments in Schizophrenia
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 540)
AUTHORS Kim,H., Shin,K. and Yi,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="SW-17"
/db_xref="taxon:9606"
/note="Schizophrenia"
gene 1. .540
/gene="env"
repeat_region 1. .540
/note="human endogenous retrovirus envelope fragment"
/rpt_family="HERV-W"

Query Match 6.4%; Score 482.4; DB 1; Length 540;
Best Local Similarity 92.8%; Pred. No. 0;
Matches 501; Conservative 4; Mismatches 35; Indels 0; Gaps 0;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||| ||| ||||||||||||||||||||||||||||||||||||| |   |||
Db         61 CTGGAAATAATCCACACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGCAACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||| |||| |||||||||| | |||||||||||||| |||||
Db        121 GCCAACTCCCAATGCATCAGGCGGGTTACTCCTCCCACGCGAATAGTCTGCCTAACCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || ||||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  |||||||||||||||||||| ||||||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||| | |||| || ||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAGCAGGTCACCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||| | |||||||||||||||||||||||||||||||||||||| :|||| ||||||||||
Db        481 ACCTAGGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTTAAAATTGAAGAGCTTTA 540
```

RESULT 638
AB115162
LOCUS AB115162 540 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:CWE15-3.
ACCESSION AB115162
VERSION AB115162.1 GI:41016597
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 540)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWE15-3"
/db_xref="taxon:9606"

```
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .540
                     /note="HERV-W env"

  Query Match             6.4%;  Score 482.4;  DB 1;  Length 540;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||| ||| |||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTTTTGCTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||| |||||||||||||||||||||||||||  | |||||
Db         61 CTGGAAATAACCCATACCTCACACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| |||||||||||||||||  |||||||||||||||| ||
Db        121 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACAGGAATAGTCTGCCTACCCCCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| ||||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        241 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| |||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| |||||||||||||||||| ||||||||||
Db        361 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        421 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||| |||||||||||||||||:|||| ||||||||||
Db        481 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATTGAAGAGCTTTA 540


RESULT 639
CQ101066/c
LOCUS       CQ101066                549 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9925 from Patent WO0157272.
ACCESSION   CQ101066
VERSION     CQ101066.1  GI:41070092
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 9925 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN PLACENTA, SIGNAL = 2.6"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  |||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||||: |||||||||||:|||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:|||| ||| |||||||||||||||||||:| |||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: ||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:|||||||||||||||:|||:|
Db         70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db         10 AGGCACTAGA 1


RESULT 640
CQ140060/c
LOCUS       CQ140060                549 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10082 from Patent WO0157276.
ACCESSION   CQ140060
VERSION     CQ140060.1  GI:41097432
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 10082 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.5"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;
```

```
Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
Db          10 AGGCACTAGA 1
```

```
RESULT 641
CQ176211/c
LOCUS       CQ176211                549 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 7607 from Patent WO0157274.
ACCESSION   CQ176211
VERSION     CQ176211.1  GI:41170950
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 7607 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN HEART, SIGNAL = 1.3"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
Db          10 AGGCACTAGA 1
```

```
RESULT 642
CQ223406/c
LOCUS       CQ223406                549 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10245 from Patent WO0157273.
ACCESSION   CQ223406
VERSION     CQ223406.1  GI:41205731
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 10245 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.2"
```

```
  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  |||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: |||||||||||:|||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| |||||||||||||||||||:| ||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:|||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1


RESULT 643
CQ261379/c
LOCUS       CQ261379                549 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9640 from Patent WO0157277.
ACCESSION   CQ261379
VERSION     CQ261379.1  GI:41233859
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 9640 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.5"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
             |||||||||::||||::| ||||:||||||||:|||||||||||||||| ||||:|:|||
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
             |||||||:|||||||||  |||||||||||::|||||||||||||  ||||||||||||
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
             |||:||||||||||  |||||||||||||||||:|||||::|||||||||||||||||||
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
             |||||||||||||||||||||: |||||||||||:|||||||||||||||||||:| ||
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  ||| |||||||||||||||||||:| ||||||||||||||||||
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||| |: | |||||||||||| ||||| |||||||||:|||
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||| ||||||:||||:|||||||||||||||:|||:|
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
              |||||||||
Db          10 AGGCACTAGA 1


RESULT 644
CQ298746/c
LOCUS       CQ298746                549 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9851 from Patent WO0186003.
ACCESSION   CQ298746
VERSION     CQ298746.1  GI:41259323
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 9851 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN LUNG, SIGNAL = 2"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;
```

```
Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
Db          10 AGGCACTAGA 1
```

```
RESULT 645
CQ335530/c
LOCUS       CQ335530                549 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9624 from Patent WO0157275.
ACCESSION   CQ335530
VERSION     CQ335530.1  GI:41284601
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 9624 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN BRAIN, SIGNAL = 1.5"

  Query Match             6.4%;  Score 481.1;  DB 1;  Length 549;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  488;  Conservative   33;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
```

```
Db        549 CTTGCAACCTGTGGCATATCTGAGTAAGGAAATTGATGTAGTGGCAAAGAGTTGGCCTCA 490

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
Db        489 TTGTTTATGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCCATTAAAATAATACA 430

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
Db        429 GGGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGG 370

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
Db        369 AGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACC 310

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
Db        309 AGTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 250

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
Db        249 AGAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGA 190

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
Db        189 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 131

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
Db        130 TTGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTGGTGAGTGACAATGGAATACTTG 71

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
Db          70 AAAGTAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 11

Qy       4047 GGGCACTAGA 4056
Db          10 AGGCACTAGA 1
```

```
RESULT 646
AB115096
LOCUS       AB115096                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE1-3.
ACCESSION   AB115096
VERSION     AB115096.1  GI:41016531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE1-3"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W env"
     repeat_region   1. .537
```

```
                     /note="HERV-W env"

  Query Match             6.4%;  Score 480.7;  DB 1;  Length 537;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   27;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             || ||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        358 GCGGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 537


RESULT 647
AB115146
LOCUS       AB115146                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE12-2.
ACCESSION   AB115146
VERSION     AB115146.1  GI:41016581
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE12-2"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 480.7;  DB 1;  Length 537;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   27;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||| ||| ||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||| |||:|||:|||||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        358 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 537


RESULT 648
AB115157
LOCUS       AB115157                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE14-3.
ACCESSION   AB115157
VERSION     AB115157.1  GI:41016592
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
```

```
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE14-3"
                     /db_xref="taxon:9606"
                     /cell_line="LOX-IMVI"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 480.7;  DB 1;  Length 537;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   27;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || |||||||| |||||||||||||||||||||||||| |||||| |||
Db        298 AATCATGTCGTACCTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTTATGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | ||| ||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||  |||||||||| ||||| ||||||||| |||| ||||||||||||||
Db        418 TACTACAAACTGCCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| ||||||| |||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 649
AB114986
LOCUS       AB114986                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-11.
ACCESSION   AB114986
VERSION     AB114986.1  GI:41016494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES            Location/Qualifiers
     source         1. .539
                    /organism="Homo sapiens"
                    /mol_type="other RNA"
                    /isolate="WE10-11"
                    /db_xref="taxon:9606"
                    /tissue_type="spleen"
     repeat_region  1. .539
                    /note="HERV-W env"

  Query Match             6.4%;  Score 480.5;  DB 1;  Length 539;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||.|||||||||||||||||| ||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||  |||||||||||||||||:|||:|||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||| ||| |||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | |||||||||||| |||||||||||||||| |||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| ||||||| |||||||||| |||||||||| |||||||||||||||
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||  |||||||||||||||| ||||||||||||||||||:||||| ||||| ||||
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 650
AB114997
LOCUS       AB114997                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-12.
ACCESSION   AB114997
VERSION     AB114997.1  GI:41016505
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
```

```
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-12"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.4%;  Score 480.5;  DB 1;  Length 539;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||| |||||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | | |||||||||||| |||||||||||||||| ||||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| ||||||| ||||||| |||||||||| |||||||| ||||||||||
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||  |||||||||||||||| |||||||||||||||||||:||||| ||||| ||||
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 651
AB115094 '
LOCUS       AB115094                 536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE1-1.
ACCESSION   AB115094
VERSION     AB115094.1  GI:41016529
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE1-1"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.4%;  Score 479.2;  DB 1;  Length 536;
  Best Local Similarity   94.8%;  Pred. No. 0;
  Matches  512;  Conservative    3;  Mismatches   21;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | ||||||  || ||||||||| |||||||||||||| |||||||||||||||||| |||
Db        297 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||
Db        357 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 536


RESULT 652
AB114988
LOCUS       AB114988                 539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-13.
ACCESSION   AB114988
VERSION     AB114988.1  GI:41016496
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-13"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 539;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   30;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| ||||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| |||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | |||||||||||| |||||||||||||||||| ||||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||||| |||||||||||||||||||:|||||  |||||  ||||
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 653
AB114995
LOCUS       AB114995                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-7.
ACCESSION   AB114995
VERSION     AB114995.1  GI:41016503
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-7"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.4%;  Score 478.9;  DB 1;  Length 539;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   30;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||| ||||| ||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACT-CTCCCCCACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||| ||||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        240 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| |||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| | | |||||||||||| |||||||||||||||||| ||||||||||
Db        360 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        420 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  |||||||||||||||||| |||||||||||||||||||:|||||  |||||  ||||
Db        480 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 539


RESULT 654
AB096028
LOCUS       AB096028                541 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-1.
ACCESSION   AB096028
VERSION     AB096028.1  GI:30698504
```

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-1"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .541
                     /gene="env"
     repeat_region   1. .541
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.4%;  Score 478.3;  DB 1;  Length 541;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCCCAGAAATGAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATATCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAAA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACAAACTCCCAATGCATCAGGTGGGTAACACCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        301 AATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        361 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACAT-GGAACGGGTCGCCGACTCCCTGGT 6677
Db        421 TACTACAAACTATCTCAAGAAATAAATGGTGACATGGGAACAGGTCACTGACTTCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
Db        481 CACCTTGCAAGATCAACTTAACTTCCCAGCAGCAGTAGTCCTTCAAAATCGAAAAGCTTT 540

Qy       6738 A 6738
```

```
Db        541 A 541


RESULT 655
AB115159
LOCUS       AB115159                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE14-5.
ACCESSION   AB115159
VERSION     AB115159.1  GI:41016594
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE14-5"
                     /db_xref="taxon:9606"
                     /cell_line="LOX-IMVI"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 477.9;  DB 1;  Length 537;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    3;  Mismatches   29;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        358 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
```

```
Db        418 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||  |||||||||||||||| |||||||||||||||||||:||||| |||||  ||||
Db        478 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 537


RESULT 656
CQ746475
LOCUS       CQ746475                660 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 32409 from Patent WO02068579.
ACCESSION   CQ746475
VERSION     CQ746475.1  GI:42367392
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 32409 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             6.4%;  Score 477.8;  DB 1;  Length 660;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  599;  Conservative    8;  Mismatches   52;  Indels   12;  Gaps    9;

Qy       1191 ATGGAGAGATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCA 1250
              |||||||  |||||  |||| |||||||||||||||||||| |||||||||||  |||||||
Db          1 ATGGAGATATATA-ATGTTACTGCTAAATCAGACACTAA-CCCAAATGAGAAAAGTGCCG 58

Qy       1251 CCATAACTGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANG 1310
              |  |||||||||||  |||:|  |||||  |||||||||||||||||||||||||||  |
Db         59 CTGTAACTGCAGCCCGAGAGATTGGCATTCTCTGGTATCTCAGTCAGGTCAATGG--TAG 116

Qy       1311 GATGACAACAGAAGGAAAGANAATGATTCCCCACAGGCCAGCARGCAGTTCCCAGTCTAS 1370
              ||||||||||| ||||||||  || ||||||||||||||  |||:|||||||||||| ||:
Db        117 GATGACAACAG-AGGAAAGAGAACGATTCCCCACAGGCACGCAGGCAGTTCCCAGTGTAG 175

Qy       1371 ACCCTCATTGGGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTA 1430
              || |||||||||   |||  |||      ||||| |||||||||||| ||||||||||||||
Db        176 ACTCTCATTGGG---ACACAAAATCAGAACAT-GGAGATTGGTGCCGCAGACATTTGCTA 231

Qy       1431 ACTTGTGTGCTASAAGGACTAAGGAAAACTASGAAGAAARTCTAYGAATTACTCAATGAT 1490
              |||||  ||||||:||||||||||||||||:|||| ||: |||:|||||||||||||||
Db        232 ACTTGCGTGCTAGAAGGACTAAGGAAAACTAGGAAG-AAGCCTATGAATTACTCAATGAT 290

Qy       1491 GTCCACCATAACACAGGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 1550
              ||||||  ||||||||  |||||  ||||||||||||||||||||||||||||||||||||
Db        291 GTCCACTATAACACA-GGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGG 349

Qy       1551 AGGCATTGAGGAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAA 1610
              ||||||||||  ||||  |  ||| ||||||||||||||||  ||||||||||||||||||
Db        350 AGGCATTGAGAAAGCATACCTGTCTGTCACCTGACTCTGTTGAAGGCCAACTAATCTTAA 409

Qy       1611 AGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCG 1669
              ||  ||||||||||||||||||||||||||||||||  |||||||||||||||  ||||
Db        410 AGGATAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAAAACTTCAAAAGTTTGCCT 469

Qy       1670 TAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAATA 1729
              |||||||  || |||||||||||||||||  |||||||||||:|||||||||||||||||
Db        470 TAGGCCCAGAACAAAACTTAGAAACCCTACTGAACTTGGCAACCTCGGTTTTTTATAATA 529

Qy       1730 GAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTA 1789
              |||||||||||||||||||||||  ||||||||||||  |||||||||||||||  ||| |
Db        530 GAGATCAGGAGGAGCAGGCGGAATGGGACAAACGGGATAAAAAAAAAGGCCACCACTTCA 589

Qy       1790 GTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 1849
              ||||||  ||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        590 GTCATGGCCCTCAGGCAAGCAGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAT 649

Qy       1850 TGAATGCCTAA 1860
              |||||||||
Db        650 CAAATGCCTAA 660


RESULT 657
AB115140
LOCUS       AB115140                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE11-1.
ACCESSION   AB115140
VERSION     AB115140.1  GI:41016575
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE11-1"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.4%;  Score 477.7;  DB 1;  Length 539;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    3;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||  ||||:||| |||||||||||| ||||||||||||
Db        240 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGCACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```

```
Db        300 AATCATGTCGTACCAAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        360 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCGGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        420 TACTACGAACTATCTCAAGAACTAAATGGTGACATGGAACAGGTCACCGACTCCCTAGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCTGTAGTAGTCCTTCAAAACCAACGAGCTTTA 539


RESULT 658
AB115108
LOCUS       AB115108                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE3-5.
ACCESSION   AB115108
VERSION     AB115108.1  GI:41016543
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE3-5"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.4%;  Score 477.6;  DB 1;  Length 536;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   22;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        178 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        297 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        357 GCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        417 TACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 536


RESULT 659
AB096027
LOCUS       AB096027                547 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-7.
ACCESSION   AB096027
VERSION     AB096027.1  GI:30698503
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hong,K., Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of HERV-W env Fragments in Fetal Brain cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 547)
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-7"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .547
                     /gene="env"
     repeat_region   1. .547
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             6.4%;  Score 477.6;  DB 1;  Length 547;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  515;  Conservative    4;  Mismatches   13;  Indels    4;  Gaps    3;

Qy       6209 GCACAGAAATAAACACC--ACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAA- 6265

Db         12 GCACAGAAATAAACCCCCACTTCCGTTTTAAGTAGGACTTCTTGTTTCCAATCTGGAAAA 71

Qy       6266 TAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACT 6325

Db         72 TAACCCATCCCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACT 131

Qy       6326 CCCAATGCATCA-GGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATA 6384

Db        132 CCCAATGCATCAGGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATA 191

Qy       6385 TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGC 6444
```

```
Db        192 TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGC 251

Qy       6445 TTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTAT 6504

Db        252 TTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTAT 311

Qy       6505 GTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGA 6564

Db        312 GTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGA 371

Qy       6565 GTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTAC 6624

Db        372 GTGCTGGGTGTACTAGGTACTGGCATTGGCGGCATCACAACCTCTACTCAGTTCTACTAC 431

Qy       6625 AAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTG 6684

Db        432 AAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTG 491

Qy       6685 CAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTAGA 6740

Db        492 CAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAAA 547


RESULT 660
AB126140
LOCUS       AB126140                538 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE2-8.
ACCESSION   AB126140
VERSION     AB126140.1  GI:47169740
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 538)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE2-8"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
                     /note="HERV-W env"
     repeat_region   1. .538
                     /note="HERV-W env"

  Query Match             6.4%;  Score 477;  DB 1;  Length 538;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  510;  Conservative    4;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACGGAAATAAACTCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 GTGGAAATAGCCCATACCTCAAACCTCACCTATGTAAAATGTAGCAATACTACAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTTGCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGTCTATGGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        300 AATTATGTCATACATAAGCCCCGCAACAAGAGAGTACCCATTCTTCCTTTTGTTATCAGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        360 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        420 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        480 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 538


RESULT 661
AB115137
LOCUS       AB115137                539 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE10-3.
ACCESSION   AB115137
VERSION     AB115137.1  GI:41016572
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE10-3"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.4%;  Score 476.4;  DB 1;  Length 539;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  511;  Conservative    3;  Mismatches   24;  Indels    4;  Gaps    2;

Qy       6198 GAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAA 6257

Db          1 GAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAA 57

Qy       6258 TSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
```

```
Db         58 TCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAC 117

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTC 6377
              | ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        118 AGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTC 177

Qy       6378 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATC 237

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
              | |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 TGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATA 297

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
              || | |||||||| |||||||||||||||||||||||||||||||||||||||||||| ||
Db        298 CAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGG 357

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGC-ATTGGCGGTATCACAACCTCTACTCAGT 6616
              |||||||||||||| | | ||| ||||||| |||||||||||||||||||||||||||||
Db        358 AGCAGGAGTGCTAGGCGGAGTAGCTACTGGCAATTGGCGGTATCACAACCTCTACTCAGT 417

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
              |||||||||||||| |||||||||||||||| ||||||||| |||||||| ||  ||||||
Db        418 TCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGG 477

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        478 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 537

Qy       6737 TA 6738
              ||
Db        538 TA 539


RESULT 662
AB115174
LOCUS       AB115174                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE18-1.
ACCESSION   AB115174
VERSION     AB115174.1  GI:41016609
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE18-1"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 476.3;  DB 1;  Length 537;
  Best Local Similarity   93.3%;  Pred. No. 0;

  Matches  504;  Conservative    3;  Mismatches   30;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| ||||| ||||||||||||||||||||||||||||||||||||  ||
Db        298 AATCATGTCATACCTAAGTCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||| |||||||    | |||||||||||| ||| ||||||||||||||||||||||||
Db        358 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        418 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 537


RESULT 663
AB097588
LOCUS       AB097588                541 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-13.
ACCESSION   AB097588
VERSION     AB097588.1  GI:31442178
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-13"
                     /db_xref="taxon:9542"
     repeat_region   1. .541
                     /note="HERV-W env"
                     /rpt_family="HERV-W"
```

```
  Query Match             6.4%;  Score 476;  DB 1;  Length 541;
  Best Local Similarity   92.9%;  Pred. No. 0;
  Matches  494;  Conservative    4;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||| ||||||||||||||||||  |||||||||||| |||||| | |||||
Db         61 CTGGAAATAACTCATACCTCAAACCTCACCTACGTAAAATTTAGCGATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCAGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTTTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||  || ||||||||||||||:|||:|||| ||||||||||||||||||||
Db        241 ATGTGCTTCCTTTCGTTCTTAGTGCCCCCTATGACCATTTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             |  |||||| || |||||||| ||||||||||||||||||||||||||||||||  ||
Db        301 AAATATGTCGTAACTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||   |||||||| |||| ||| ||||| ||||||||||||||||||||
Db        361 GCAGGAGTGCTAGGCAGACTAGGTATTGGCGTTGACGGTACCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||| |||||||||||||||||  |||||||||||||| || ||||||||||||
Db        421 TACTACAAACTGTCTCAAGAACTAAATGGTAACATGGAACGGGTCACCAACTCCCTGGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAA 6730
             |||||||||||||||||||||||||||||||||||| |||||:||||| ||
Db        481 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGGCCTTCAAAATCAAA 532


RESULT 664
AB114998
LOCUS       AB114998                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE11-13.
ACCESSION   AB114998
VERSION     AB114998.1  GI:41016506
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE11-13"
                     /db_xref="taxon:9606"
                     /tissue_type="thymus"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 475.9;  DB 1;  Length 537;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   30;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||||  ||||    ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||||||||||||||||||| ||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| |||||||||||||||| |||||||||
Db        358 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||  |||||||  |||||||  |||||||||  |||||||||  |||||||||
Db        418 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||  ||||||||||||||||  ||||||||||||||||||:|||||  |||||  ||||
Db        478 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 537


RESULT 665
AB115164
LOCUS       AB115164                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE15-5.
ACCESSION   AB115164
VERSION     AB115164.1  GI:41016599
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
```

```
                     /mol_type="other RNA"
                     /isolate="CWE15-5"
                     /db_xref="taxon:9606"
                     /cell_line="AZ521"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.4%;  Score 475.9;  DB 1;  Length 537;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   30;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        118 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||| |||||||||||||||| |||||||||||||||||||| |||||
Db        178 GGAATATTTTTTGTCTGTAGTACCTCAGCCTATCGTCGTTTGAATGGCTCTTCAAAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||||||||:|||:|||||||||||| ||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGCACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||||||||||||||||||||||||||| |||||||| | |||
Db        298 AATCATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTACTTTTGTTGTGGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||| | |||| | | |||||||||||||||||||||||||||||||||||||||||
Db        358 GCAGGAGGGTTAGGCGGATTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| ||||||||| |||  || |||||||| |||
Db        418 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAATGGGCTGCTGACTCCCTAGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||| |||||||| ||||:||||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGTAGCAGTAGTTCTTCAAAATCGAAGAGCTTTA 537


RESULT 666
AB126158
LOCUS       AB126158                536 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-2.
ACCESSION   AB126158
VERSION     AB126158.1  GI:47169758
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
```

```
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-2"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.4%;  Score 475.6;  DB 1;  Length 536;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    4;  Mismatches   23;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||||| |||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| |||||| || :|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||  |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        357 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||| |||||||||| |||||||||:|||| ||||||||||||||||
Db        477 ACCTTGCAAGATCAACTTAATTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 667
AB115129
LOCUS       AB115129                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE8-4.
ACCESSION   AB115129
VERSION     AB115129.1  GI:41016564
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE8-4"
                     /db_xref="taxon:9606"
                     /cell_line="Hep G2"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.3%;  Score 474.4;  DB 1;  Length 536;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches  509;  Conservative    3;  Mismatches   24;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| | |||||||||||||||||||||||||||||||||||||||||||||  ||
Db        297 AATTATGTCGTGCCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||||
Db        357 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||| ||||:||||||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 536


RESULT 668
AB115107
LOCUS       AB115107                 537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE3-4.
ACCESSION   AB115107
VERSION     AB115107.1  GI:41016542
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
    PUBMED  15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE3-4"
                     /db_xref="taxon:9606"
                     /cell_line="BT-474"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.3%;  Score 474.3;  DB 1;  Length 537;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   31;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||| ||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAGCACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCTTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||| |||||||| |||||||||| ||| ||||| ||||
Db        178 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCATTGTTTGAATAGCTTTTCAGTATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||||:|||:|||||||||||||  |||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAATGAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||||| ||||||||||||||||||||||||||||||||||||||||  ||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||   |||||||||||||||||| ||||||||||||||||||||||||||
Db        358 GCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||| ||||| ||||||||| |||| ||||| |||||||||
Db        418 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||| ||||||||| ||||:|||||||||||||||
Db        478 ACCTTGCAAGATCAACTTAACTCCCTAGTAGCAGTAGTTCTTCAAAATCGAAGAGCTTTA 537


RESULT 669
AB115250
LOCUS       AB115250                 661 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP1-4.
ACCESSION   AB115250
VERSION     AB115250.1  GI:41052478
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 661)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .661
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP1-4"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W pol"
     repeat_region   1. .661
                     /note="HERV-W pol"

  Query Match             6.3%;  Score 473.5;  DB 1;  Length 661;
  Best Local Similarity   83.9%;  Pred. No. 0;
  Matches  557;  Conservative   58;  Mismatches   38;  Indels   11;  Gaps    9;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCC 3149
Db          1 TCCAAACCAAAAGTTCAGCTCTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATCC 60

Qy       3150 AAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYC 3209
Db         61 AAAGGCACCAGGGCCCTCAGTGAGGAACATATCCAGCCTATACTGGCTTATCCTCATCCC 120

Qy       3210 AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATT 3269
Db        121 AAAACGCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGTCGAATATGGATT 180

Qy       3270 CCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGC 3329
Db        181 -CCCAGGTACGGTGAAATAGACAGACCATTGTATACACTAATTAAGGAAACTTGGAAAGT 239

Qy       3330 CAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGA 3389
Db        240 CAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAA 299

Qy       3390 AGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAY 3449
Db        300 GGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCGATGGGGCAAGACTTTTCTTTATAT 357

Qy       3450 RTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTT 3509
Db        358 GTCATA-AAAAAACCAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACGAGCTT 416

Qy       3510 GCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTG 3569
Db        417 GCAACCCGTGGCATACCTGAATAAGGAAATTGATGTACTGGCAAAGGGTTGGCCTCATTG 476

Qy       3570 TTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAA-ATAATACAGG 3628
Db        477 TTTACGGGTAGTGGCAACAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGGTAATACAGG 536

Qy       3629 G-RAGAGATCTTACTGTGTGGACATCTCATGAKGTGAA--YRGCATACTCACTGCTAAAG 3685
Db        537 GAAAGAGATCTTACTGGGTGAACATCTCATGATGTGAAATGGGCATACTCCCTGCTAAAG 596

Qy       3686 GAGACTTGTGGC-TGTCAGACAA-CYGTTTACTTAAATRTCAGGCTC-TATTACTTGAAR 3742
Db        597 GAGACTTGGGGCTTGTCAGACAACCCGGTTACTTAAATATCAGGCTCTTATTACTTGATG 656

Qy       3743 GGCC 3746
Db        657 GGGC 660


RESULT 670
AB126152
LOCUS       AB126152                538 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-2.
ACCESSION   AB126152
VERSION     AB126152.1  GI:47169752
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 538)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-2"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .538
                     /note="HERV-W env"

  Query Match             6.3%;  Score 473.4;  DB 1;  Length 538;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  508;  Conservative    4;  Mismatches   25;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCTTACCTAAGCCCCGCAACAGAAGAGTATCCATTCTTCCTTTTGTTATCAGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        358 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        418 TACTACGAACTATCTCAAGAGCTAAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 477
```

```
Qy       6679 ACCTTGCAAGATC-AACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             ||||||||||||||| |||||||||||||||| ||| ||||||||||:||||| |||||||||
Db        478 ACCTTGCAAGATCAAACTTAACTCCCTAGAAGCTGTAGTCCTTCGAAATCAAAGAGCTTT 537

Qy       6738 A 6738
              |
Db        538 A 538
```

RESULT 671
AB114929
LOCUS AB114929 537 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE1-18.
ACCESSION AB114929
VERSION AB114929.1 GI:41016437
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 537)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .537
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE1-18"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region 1. .537
/note="HERV-W env"

Query Match 6.3%; Score 473.1; DB 1; Length 537;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 502; Conservative 3; Mismatches 32; Indels 3; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| |||| ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

```
Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | | |||||||||||| |||||||||||||||||| |||||||||
Db        358 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGCACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        418 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              || ||||||||||||||||| |||||||||||||||||||:||||| | |||||||
Db        478 GCCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTTTA 537
```

RESULT 672
AB114993
LOCUS AB114993 537 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:WE11-4.
ACCESSION AB114993
VERSION AB114993.1 GI:41016501
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 537)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .537
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="WE11-4"
/db_xref="taxon:9606"
/tissue_type="thymus"
repeat_region 1. .537
/note="HERV-W env"

Query Match 6.3%; Score 473.1; DB 1; Length 537;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 502; Conservative 3; Mismatches 32; Indels 3; Gaps 1;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||| |||||||||| |||| ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 GTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTCATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
```
```

```
Db        298 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        358 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        418 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        478 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 537


RESULT 673
AB115171
LOCUS       AB115171                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE17-3.
ACCESSION   AB115171
VERSION     AB115171.1  GI:41016606
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE17-3"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.3%;  Score 472.8;  DB 1;  Length 536;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   25;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498

Db        238 GTGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558

Db        297 AATTATGTTGTACCTAAGCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618

Db        357 GCAGGAGTGCTAGGTGGACTAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678

Db        417 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738

Db        477 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 674
AB079505
LOCUS       AB079505                539 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-5.
ACCESSION   AB079505
VERSION     AB079505.1  GI:22122079
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-5"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .539
                     /rpt_family="HERV-W"

  Query Match             6.3%;  Score 472.5;  DB 1;  Length 539;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   34;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258

Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318

Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTATGTAAAATTTAGCAATACTATAGATACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378

Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTTCTCCCACACAAACAGTCTGCCTACTCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438

Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240
```

```
Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        241 ATGTGCTTCCTCTCATTCTTAATG-CCCCTATGACCATCTACACTGAACAAGATTTATGC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        300 AATCATGTTGTACCTAAGCCCTGCAACAAAAGAGCACCCTTTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        360 GTAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGTGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        420 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAATGAGTCGCTGACTCTCTAATC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        480 ACCTTGCAAGATCAACTTAATTCCCTAGCAACAGTAGTCCTTTAAAATCAAAGAGCTTTA 539
```

RESULT 675
AB115252
LOCUS AB115252 656 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W pol, isolate:CWP2-1.
ACCESSION AB115252
VERSION AB115252.1 GI:41052480
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 656)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .656
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWP2-1"
/db_xref="taxon:9606"
/cell_line="PFSK-1"
/note="HERV-W pol"
repeat_region 1. .656
/note="HERV-W pol"

Query Match 6.3%; Score 472.5; DB 1; Length 656;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 550; Conservative 58; Mismatches 43; Indels 13; Gaps 8;

```
Qy       3089 TTCCAAACSARARGCTCARCTCTGCT-CACAGCAGGTTAAATAC--TTAGGRCTAARATT 3145
Db          1 TCCCAACAAAAATGCTCAGCCCTGCTCCCCAGCAGGTTAAAAACCTTAGGGGTAAAAATT 60

Qy       3146 ATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCA-GCCTATACTGGCTTATCCTC 3204
Db         61 ATCCAAAGGCCCAAGGGCCCTCAGTGAGGGACGTATCCAGGCCTATACTGGCTTATCCTC 120

Qy       3205 ATCYCAAAACCCTAAAGC-AACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWA 3263
Db        121 ATCCCAAAACCCTAAAGCAAATTAGGAGGGTTCCTTGGCATAACAAGCTTCTGCCGAATA 180

Qy       3264 TGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCA 3323
Db        181 TGGATT-CCCAGGTACAGCAAAATAGTCAGGCCATTATACACAC----TAAGGAAACTCA 235

Qy       3324 GAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCC 3383
Db        236 GAAAGCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGTGGCTTTTCGGGCCCT 295

Qy       3384 TAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTST 3443
Db        296 AAAGAAGGCCCT--AACCCAAGCCCCAGTGGTAAGCTTGCCAATGGGGCAAGACTTTTCT 353

Qy       3444 TYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAY 3503
Db        354 TTATATGTCGCA-AAAGAAACAGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGGGAC 412

Qy       3504 GAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCY 3563
Db        413 CAGCTTGCAACCTGTGGCATACCTGAGTAAGGAAATTGATGTAGTTGCAAAGGGTTGACC 472

Qy       3564 TCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAAT 3623
Db        473 TCATTGTTTACAGGTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAGTAAT 532

Qy       3624 ACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAA 3683
Db        533 ACAGGGAAGAGATCTTACTGTGTGAACATCTCATGATGTGAATGGCATACTCACTGCTAA 592

Qy       3684 AGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARG 3743
Db        593 AGGAGATTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGATGG 652

Qy       3744 GCCA 3747
Db        653 GCCA 656
```

RESULT 676
AB126147
LOCUS AB126147 536 bp RNA linear PRI 14-MAY-2004
DEFINITION Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE5-4.
ACCESSION AB126147
VERSION AB126147.1 GI:47169747
KEYWORDS .
SOURCE Macaca fuscata (Japanese macaque)
ORGANISM Macaca fuscata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS Yi,J., Takenaka,O. and Kim,H.
TITLE Expression and identification of HERV-W family in Japanese Monkeys (Macaca fuscata)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 536)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .536
/organism="Macaca fuscata"
/mol_type="other RNA"
/isolate="JM-WE5-4"
/db_xref="taxon:9542"
/tissue_type="thymus"
/note="HERV-W env"
repeat_region 1. .536
/note="HERV-W env"

```
  Query Match             6.3%;  Score 472.4;  DB 1;  Length 536;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  507;  Conservative    4;  Mismatches   25;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||| ||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||  |||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||||| ||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||||||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 536


RESULT 677
AB115125
LOCUS       AB115125                534 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE7-6.
ACCESSION   AB115125
VERSION     AB115125.1  GI:41016560
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 534)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE7-6"
                     /db_xref="taxon:9606"
                     /cell_line="Jurkat"
                     /note="HERV-W env"
     repeat_region   1. .534
                     /note="HERV-W env"

  Query Match             6.3%;  Score 471.4;  DB 1;  Length 534;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  507;  Conservative    3;  Mismatches   24;  Indels    6;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||| |||||||||| ||||   |||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||  | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||:||| |||||||||||||||| ||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAAATTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||   | ||||||||||||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | |||||||||||||  |||||||||||| | ||||||||||
Db        358 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
               |||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||
Db        417 --CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 474

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             || ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        475 ACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 534


RESULT 678
AB051002
LOCUS       AB051002                536 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-5-8.
ACCESSION   AB051002
VERSION     AB051002.1  GI:14020827
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
```

```
                     /mol_type="genomic DNA"
                     /isolate="W-5-8"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     repeat_region   1. .536
                     /rpt_family="HERV-W env"

  Query Match             6.3%;  Score 471.2;  DB 1;  Length 536;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   32;  Indels    4;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||    || ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAAT----ACTACTTCCTTTTTAGTAGGACCTCTTGTTTCCAAT 56

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||| ||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         57 CTGGAAACAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 116

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        117 ACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 176

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||| |||||||||||||||  |||||||||||||||| |||||
Db        177 GGAATATTTTTTGTCTGTAGTACCTCAGCCTATCGTCATTTGAATGGCTCTTCAAAATCT 236

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||| ||||||||||||
Db        237 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGCACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || | ||||||| |||||||||||||||||||||||||||||||  ||
Db        297 AATCATGTCGTACCAAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||| |||||||||||||||||| |||||||||||||||||| ||||
Db        357 GCAGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCGGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||||||||| |||||||||| |||| |||||||||| |||
Db        417 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACAGGTCACCGACTCCCTAGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||| | ||||||||||:||| | | ||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCTGAAGTAGTCCTTCAAAACCAACGAGCTTTA 536


RESULT 679
AB114985
LOCUS       AB114985                537 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE10-9.
ACCESSION   AB114985
VERSION     AB114985.1  GI:41016493
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)

FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE10-9"
                     /db_xref="taxon:9606"
                     /tissue_type="spleen"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.3%;  Score 471.1;  DB 1;  Length 537;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   33;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||| |||||| ||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAGTGCATCAGATGGGTAGCTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGCCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||| ||||||||||||||:|||:|||||||||||||||||||||| ||||
Db        238 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||||||  |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 357

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              || |||||||||||  | | |||||||||||| |||||||||||||||||| ||||||||||
Db        358 GCGGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 417

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| ||||||| ||||||| |||||||||| ||||||||| |||||||||
Db        418 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||  ||||||||||||||||||||||||||||||||||:||||| ||||| ||||
Db        478 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 537


RESULT 680
AB079508
LOCUS       AB079508                543 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-8.
ACCESSION   AB079508
VERSION     AB079508.1  GI:22122082
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
```

```
          of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
          Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-8"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .543
                     /rpt_family="HERV-W"

  Query Match             6.3%;  Score 470.4;  DB 1;  Length 543;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  507;  Conservative    3;  Mismatches   29;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              |||| |||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 TTGGAGATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||||||||| || ||||| |||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTCAAGGGCTCATCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             || |||||||||||||||||||||||:|||:|||||||||||||||| |||||||||
Db        241 ATATGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAGGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| || |||||||||||||||||||||||||||||||||||||||||||| |||
Db        301 AATCATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||| |||||||||| | ||||||| |||||||||| ||||||||||||||||||||||| |
Db        361 GCAAGAGTGCTAGGCGGACTAGGTTCTGGCATTGGTGGTATCACAACCTCTACTCAGTCC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAC-GGGTCGCCGACTCCCTGGT 6677
             |||||||||||||||||||||||||||| |||| ||||| ||||||| ||||||||||
Db        421 TACTACAAACTATCTCAAGAACTAAATGGTGACACGGAACGGGGTCGCTGACTCCCTGGT 480

Qy       6678 CACCTTGCAAGATCAACTTAAC-TCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             ||| ||||||||||||||||||| ||||||||||||||||||||:|||||  | |||||
Db        481 CACATTGCAAGATCAACTTAACTTCCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTT 540

Qy       6737 T 6737
              |
Db        541 T 541


RESULT 681
AB115152
LOCUS       AB115152                544 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE13-4.
ACCESSION   AB115152
VERSION     AB115152.1  GI:41016587
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 544)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE13-4"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
                     /note="HERV-W env"
     repeat_region   1. .544
                     /note="HERV-W env"

  Query Match             6.3%;  Score 470.4;  DB 1;  Length 544;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches  508;  Conservative    3;  Mismatches   29;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAA---CACCACTTCCGTTTTAGTAGGACCTCTTGTTT-C 6254
             |||||||||||||||||  ||   | || |||| |||||||||||||||||||||||| |
Db          1 AACAACTTCAGCACAGAAAATAAACCCCCCGCTTCTGTTTTAGTAGGACCTCTTGTTTCC 60

Qy       6255 CAATSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATA 6314
             ||||:|||||||||||||| |||||||||||| |||||||||||||||||||| | |
Db         61 CAATGTGGAAATAACCCATACTTCAAACCTCACCCGTGTAAAATTTAGCAATACTATAGA 120

Qy       6315 CACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 6374
             |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        121 CACAACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACC 180

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 240

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             |||||||||||||||||||||||||||||| ||||:||| ||||||||||||||||||||
Db        241 ATCTATGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTT 300

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||| | ||||| || ||||||||| ||||||||||||||||||||||||||||||||||
Db        301 ATACAATCATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 360

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
               |||||||||||||||  |||||||||||||||||||||||||||| ||||||||||||
Db        361 CAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCA 420

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCT 6674
             |||||||||||||||||||||||| ||||||| |||||||||| |||| | ||||||||
Db        421 GTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCT 480

Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        481 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGC 540

Qy       6735 TTTA 6738
              ||||
Db        541 TTTA 544


RESULT 682
AB114971
LOCUS       AB114971                535 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE9-1.
ACCESSION   AB114971
VERSION     AB114971.1  GI:41016479
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE9-1"
                     /db_xref="taxon:9606"
                     /tissue_type="skeletal muscle"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.3%;  Score 469.7;  DB 1;  Length 535;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches  494;  Conservative    4;  Mismatches   27;  Indels    1;  Gaps    1;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
             | | ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db         11 ACACATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             |||||||||||||||||||||||||||||||| | | ||||||||||||||||||||
Db         71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
             |||||||||||||||||| ||| ||| ||||||||||||||||||||||||||||||||||
Db        131 CATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 189

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             |||||||||||||||||||  |||||||||||| ||| ||||||||||||||||||||||
Db        190 CTGTGGTACCTCAGCCTATCACTGTTTGAATGGCTTTTCGGAATCTATGTGCTTCCTCTC 249

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
             |||||||||||| |||:|||:|||||||||||||||||||||||||||| ||||||| || |
Db        250 ATTCTTAGTGCACCCTATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTACC 309

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             |||||||||||||||||||||||||||||||||||||||||||  |||||||| |||||||
Db        310 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGAGCAGGAATGCTAGG 369

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
               | |||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||
Db        370 CAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 429

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             |||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        430 TCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCA 489

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||:||||||||||||||||||
Db        490 ACTTAACTCCCTAGCAGCAGTAGTACTTCAAAATCGAAGAGCTTTA 535


RESULT 683
AB115166
LOCUS       AB115166                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE16-2.
ACCESSION   AB115166
VERSION     AB115166.1  GI:41016601
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE16-2"
                     /db_xref="taxon:9606"
                     /cell_line="2F7"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.3%;  Score 469.6;  DB 1;  Length 536;
  Best Local Similarity   93.7%;  Pred. No. 0;
  Matches  506;  Conservative    3;  Mismatches   27;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        297 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||  |||||||    | |||||||||||||| ||| ||||||||||||||||||||
Db        357 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||| ||||| ||||||||| |||| ||||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 684
```

```
AB050996
LOCUS       AB050996                537 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens human endogenous retrovirus W gene for envelope
            protein, partial cds, isolate:W-1-1.
ACCESSION   AB050996
VERSION     AB050996.1  GI:14020818
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-1-1"
                     /db_xref="taxon:9606"
                     /chromosome="1"
     gene            1. .537
                     /gene="env"
     CDS             <1. .>537
                     /gene="env"
                     /codon_start=1
                     /product="envelope protein"
                     /protein_id="BAB47555.1"
                     /db_xref="GI:14020819"
                     /translation="NNFSTEINTTSILVGPLVSNLEIIHTSNLTCVKFSNTIATANSQ
                     CIRRVTPPTRIVCLTSGIFFVCGTSAYRCLNGSSESMCFLSFLVPPMTIYTEQNLYNH
                     VAPKPRNKRVPILPFVIGAGVLGGLGTGPGGITASTQFYKLSQELNGDMERVADSLVT
                     LQDQLNSLAAVVLQNQTAL"
     repeat_region   1. .537
                     /rpt_family="HERV-W"

  Query Match             6.3%;  Score 469.5;  DB 1;  Length 537;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    4;  Mismatches   34;  Indels    3;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             || |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||| |||||||||||||||||||||||||||||||||||||| |    |||
Db         61 CTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGCAACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||| |||| |||||||||| | |||||||||||||| |||||
Db        121 GCCAACTCCCAATGCATCAGGCGGGTTACTCCTCCCACGCGAATAGTCTGCCTAACCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAAATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||   | |||||||||||||||||||||||||||||||||||||||||||| |||

Db        301 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | |||||||||||||  |||||||||||| | ||||||||||
Db        361 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        420 --CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 477

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              || |||||||||||||||||||||||||||||||||||||||:|||||  | |||||||
Db        478 ACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTTTA 537


RESULT 685
AB126151
LOCUS       AB126151                539 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-1.
ACCESSION   AB126151
VERSION     AB126151.1  GI:47169751
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-1"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             6.3%;  Score 469.3;  DB 1;  Length 539;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    4;  Mismatches   36;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||| |||||||||||||||||| | | |||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTATGTAAAATTTAGCAATACTATAGATACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||| |||||||||| |||||||||| ||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTTCTCCCACACAAACAGTCTGCCTACTCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGGATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| || ||||:|||:|||||||||||||||||||||||| |
Db        241 ATGTGCTTCCTCTCATTCTTAATG-CCCCTATGACCATCTACACTGAACAAGATTTATGC 299
```

```
Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        300 AATCATGTTGTACCTAAGCCCTGCAACAAAAGAGCACCCTTTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        360 GTAGGAGTGCTAGGCAGACTAGGTACTGGCATTGATGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        420 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAATGAGTCGCTGACTCTCTAATC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        480 ACCTTGCAAGATCAACTTAATTCCCTAGCAACAGTAGTCCTTTAAAATCAAAGAGCTTTA 539


RESULT 686
AB126159
LOCUS       AB126159                536 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-4.
ACCESSION   AB126159
VERSION     AB126159.1  GI:47169759
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-4"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.3%;  Score 469.2;  DB 1;  Length 536;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   27;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---CTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
Db        417 TACTACGAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 536


RESULT 687
AB126165
LOCUS       AB126165                536 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE9-6.
ACCESSION   AB126165
VERSION     AB126165.1  GI:47169765
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE9-6"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.3%;  Score 469.2;  DB 1;  Length 536;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   27;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
```

```
                |||||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||||||||||||| |||||| ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||| |||||||||||||||||| |||||||||||||| ||||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||||||||||||||||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 536


RESULT 688
AB115251
LOCUS       AB115251                652 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP1-5.
ACCESSION   AB115251
VERSION     AB115251.1  GI:41052479
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 652)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP1-5"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
                     /note="HERV-W pol"
     repeat_region   1. .652
                     /note="HERV-W pol"

  Query Match             6.3%;  Score 468.6;  DB 1;  Length 652;
  Best Local Similarity   82.5%;  Pred. No. 0;
  Matches  546;  Conservative   59;  Mismatches   43;  Indels   14;  Gaps    8;

Qy       3090 TCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTA-TC 3148
                |||||||:|: :| |||: | || ||| |||||||||||||| ||||:  ||:|||| ||
Db          1 TCCAAACCAAGGGTTCAGTTTTGGTCATAGCAGGTTAAATACGTAGGGGCAAAATTATTC 60

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAY-RYATCCAGCCTATACTGGCTTATCCTCATC 3207
                ||||||| |||:||| |||||||||| |: ::|||||||||||||||||||||||||||
Db         61 CAAAGGCCCCAGGGCACTCAGTGAGGGACGGTATCCAGCCTATACTGGCTTATCCTCATC 120

Qy       3208 YC--AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATG 3265
                :|  ||| |||||||||| | | : ::|||||||||:|||:|||:||||||||||:|||
Db        121 CCCAAAACCCCTAAAGCAAATTAAGAGGTTCCTTGGCATAACAGGTTTCTGCCGAATATG 180

Qy       3266 GATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGA 3325
                |||| || ||||:||||:|||||||||| :|||||:|||||:|||||||||||||||||
Db        181 GATT-CCTAGGTATGGCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGA 239

Qy       3326 AAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTA 3385
                || |||||||||||||:|||||||||:|:||||| :::|||||||||||||||||||  |
Db        240 AATCCAATACCCATTTAGTAAGATGGACACCTGAATCAGAAGTGGCTTTCCAGGCCCTAA 299

Qy       3386 AAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTY 3445
                |  | | |||  |||| |||:|||||||||||:||| ||| :||||||||||  :|||:
Db        300 AGAAGGCCCT--AACCGAAGCCCCAGTGTTAAGCTTGTCAATGGGGCAAGACTTGACTTT 357

Qy       3446 ATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGA 3505
                |||::|||||||||| |||||:||:||||||:|||||||||||||||:|||:|||||: |
Db        358 ATATGTCACAGAAAAGAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGACAA 417

Qy       3506 GCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTC 3565
                ||||||||||::||||::||||||:||||||||:|||||||||||| ||||||||:|:||
Db        418 GCTTGCAACCTGTGGCATACCTGAGTAAGGAAACTGATGTAGTGGC-AAGGGTTGGCCTC 476

Qy       3566 ATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATAC 3625
                || |||||: |||||||   |||||||||| ::|||||||||||||||||||||||||||
Db        477 ATGGTTTACAGGTAGTGACAGCAGTAGCAGGCTTAGTATCTGAAGCAGTTAAAATAATAC 536

Qy       3626 AGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAG 3685
                |||     |||||||||||||||||||| |||||:||||||::|||||||  ||||||||||
Db        537 AGG-----GATCTTACTGTGTGGACATC-CATGATGTGAATGGCATACTAACTGCTAAAG 590

Qy       3686 GAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGC 3745
                |||||||||||||||||||||: |||||||||||:|| ||||||||||||||:|||
Db        591 GAGACTTGTGGCTGTCAGACAACCATTTACTTAAATATCGGGCTCTATTACTTGAAGGGC 650

Qy       3746 CA 3747
                ||
Db        651 CA 652


RESULT 689
AB126153
LOCUS       AB126153                540 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-3.
ACCESSION   AB126153
VERSION     AB126153.1  GI:47169753
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-3"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .540
```

```
                    /note="HERV-W env"

  Query Match             6.3%;  Score 468.4;  DB 1;  Length 540;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  505;  Conservative    4;  Mismatches   30;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACGGAAATAAACTCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||| |||||||||||||||||||||| |||||||| |||||||||||||| |||||
Db         61 GTGGAAATAGCCCATACCTCAAACCTCACCTATGTAAAATGTAGCAATACTACAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||  ||||| ||||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTTGCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||| |||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGTCTATGGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || ||||||||||||||| ||||||| |||||||||||||||||||| |||
Db        300 AATCATGTCTTACCTAAGCCCCGCAACAGAAGAGTATCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||| | |||||||| ||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTATTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||| |||||||||||||| |||||||| ||||||||| |||| ||||||||||||||||
Db        420 TACTACGAACTATCTCAAGAGCTAAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATC-AACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              |||||||||||||| |||||||||||||| ||| ||||||||||:||||| |||||||||
Db        480 ACCTTGCAAGATCAAACTTAACTCCCTAGAAGCTGTAGTCCTTCGAAATCAAAGAGCTTT 539

Qy       6738 A 6738
              |
Db        540 A 540


RESULT 690
AB097585
LOCUS       AB097585                539 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-10.
ACCESSION   AB097585
VERSION     AB097585.1  GI:31442175
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-10"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.2%;  Score 468.1;  DB 1;  Length 539;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    3;  Mismatches   37;  Indels    1;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||| |||||| ||| |||||| |||| ||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACTACTTCCATTTGAGTAGGGCCTCATGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||| || ||||||||| |||||||||| |||| |    ||||
Db         61 CTGGAAATAACCCATACCTCCAATCTCACCTGTATAAAATTTAGCCATACGATTGGCACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              | |||||||  |||| |||||||||| |||||||||||| |||||||||||||||||||
Db        121 ATCAACTCCCGGTGCAGCAGGTGGGTGACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTGGTGTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||||||||||||||||||||||||||||| |||:|||||||||||||||||||||| ||
Db        241 GTGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATCTACACTGAACAAGATTTAAAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              ||||||||| || |||||||| |||||||||||||||||||||||||||||||||| |||
Db        300 AGTTATGTCGTACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||||  ||||||||| |||| | ||||||||||||
Db        420 TACTACAAACTGTCTCAAGAACTAAATGGTAACATGGAACAGGTCACTGACTCCCTGGTC 479

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||||:|||| ||||||||||
Db        480 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATTGAAGAGCTTTA 539


RESULT 691
AB115177
LOCUS       AB115177                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE18-4.
ACCESSION   AB115177
VERSION     AB115177.1  GI:41016612
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
```

```
           of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
           Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE18-4"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.2%;  Score 468;  DB 1;  Length 536;
  Best Local Similarity   93.5%;  Pred. No. 0;
  Matches  505;  Conservative    3;  Mismatches   28;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db        118 ACCAACTCCCAGTGCATCAGATGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||| |||:|||| ||||||||||||||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | ||||||| || ||||||||||||||||||||||||||||||||||||||||||  ||
Db        297 AATTATGTCGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCCGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||| |||||||   | |||||||||||||| ||| ||||||||||||||||||||||
Db        357 GCAGGAATGCTAGGCAGATTAGGTACTGGCATGGGCAGTATCACAACCTCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||| | ||| |||||||||| |||| ||||||||||||||
Db        417 TACTACAAACTATCTCAAGAACTCAGTGGTGACATGGAATGGGTTGCCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 692
AB115178
LOCUS       AB115178                 535 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:CWE18-5.
ACCESSION   AB115178
VERSION     AB115178.1  GI:41016613
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWE18-5"
                     /db_xref="taxon:9606"
                     /cell_line="C-33A"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.2%;  Score 467.6;  DB 1;  Length 535;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  505;  Conservative    3;  Mismatches   27;  Indels    6;  Gaps    2;

Qy       6198 GAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAA 6257
             |||||||||||||||||||||||||||||||| |||||||||| ||||   |||||||
Db          1 GAACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAA 57

Qy       6258 TSTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
              |:|||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db         58 TCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACAC 117

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTC 6377
             | |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        118 AGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTC 177

Qy       6378 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        178 AGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATC 237

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             | |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||||
Db        238 TGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATA 297

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             || | ||||   | ||||||||||||||||||||||||||||||||||||||||||| ||
Db        298 CAATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGG 357

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
             |||||||||||||| | |||||||||||| ||||||||||| | ||||||||||||||||
Db        358 AGCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT 417

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
                |||||||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        418 ---CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGT 474

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             ||| |||||||| ||||||||||||||||||||||||||||:|||||||||||||||||
Db        475 CACATTGCAAGACCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 534

Qy       6738 A 6738
              |
Db        535 A 535


RESULT 693
AB126161
LOCUS       AB126161                 536 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-8.
ACCESSION   AB126161
VERSION     AB126161.1  GI:47169761
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-8"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.2%;  Score 467.6;  DB 1;  Length 536;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  504;  Conservative    4;  Mismatches   28;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||| |||| ||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||||: ||:||||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  || |||||||||||||| |||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||| |||||||||||||||||||| |||||||||||||| | ||||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGTCGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 536


RESULT 694
AB115265
LOCUS       AB115265                644 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W pol, isolate:CWP5-3.
ACCESSION   AB115265
VERSION     AB115265.1  GI:41052493
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 644)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-51510-3346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="CWP5-3"
                     /db_xref="taxon:9606"
                     /cell_line="TE-1"
                     /note="HERV-W pol"
     repeat_region   1. .644
                     /note="HERV-W pol"

  Query Match             6.2%;  Score 467.4;  DB 1;  Length 644;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches  553;  Conservative   56;  Mismatches   32;  Indels   20;  Gaps    9;

Qy       3090 TCCAAACSARARGCTCAR-CTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATC 3148
              |||||||:|:|:| |||: || ||||| ||||||||| ||||||||||:||||:||||||
Db          1 TCCAAACCAAAAGTTCAGGCTTTGCTCTCAGCAGGTTTAATACTTAGGGCTAAAATTATC 60

Qy       3149 CAAAGGCACCARGGCCCTCAGTGAGGAAYRYAT-CCAGCCTATACTGGCTTATCCTCATC 3207
              ||||||||||:|||||||||||||||:::|| |||||||||||||||||||||||||||||
Db         61 CAAAGGCACCAGGGCCCTCAGTGAGGAACATATCCCAGCCTATACTGGCTTATCCTCATC 120

Qy       3208 YC-AAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGG 3266
              :| |||||||||||||||||| :|::|||||      |||: ||:||||||||||:||||
Db        121 CCAAAAACCCTAAAGCAACTAATAGGGTTCCT-----TAACGGGTTTCTGCCGAATATGG 175

Qy       3267 ATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAA 3326
              ||| |||||||: || :|||||  ||||  |||| :|||||:||||||||||||||||||
Db        176 ATT-CCCAGGTACGGTGAAATACCCAGAACATTCTATACACTAATTAAGGAAACTCAGAA 234

Qy       3327 AGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAA 3386
              |||||||||||||||||:||||||||:|:|||||| :|||| |||||| ||| ||  |
Db        235 AGCCAATACCCATTTAGTAAGATGGACACCTGAAGCGGAAGTAGCTTTCTAGGACCTAAC 294

Qy       3387 AGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYA 3446
              | | |||  ||||||||:|||||||||||||:|||||||||:|| ||||| ||||:||:|
Db        295 GAAGGCCCT--AACCCAAGTCCCAGTGTTAAGCTTGCCAACAGGACAAGATTTTTCTTTA 352

Qy       3447 TAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAG 3506
              ||::|||||| |||||||||:||:|||||||:|||||||||||||||:|||:|||||:|||
Db        353 TATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGGTCCGAGGGATGAG 411

Qy       3507 CTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCA 3566
              |||||||||::||| ::|||||:||||||||:|||||||||| |||||||  |:|||
Db        412 CTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGTAAAGGGTTAGCCTCA 471

Qy       3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACA 3626
              |||||||: |||||||||||||||||||||::|||||||||||||||||||||||||| |||
Db        472 TTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAACACA 531

Qy       3627 GGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGG 3686
              |||:|||||||||||||||||      ||||| ||||||::|||||||||||||||||||
Db        532 GGGAAGAGATCTTACTGTGTGGA-----CATGACGTGAACGGCATACTCACTGCTAAAGA 586

Qy       3687 AGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCC 3746
```

```
              ||||||   |||||||||||||:||||||||||||:||||||||||||||||||||:||||
Db        587 AGACTT---GCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCC 643

Qy       3747 A 3747
              |
Db        644 A 644
```

RESULT 695
AB097584
LOCUS AB097584 539 bp DNA linear PRI 05-AUG-2003
DEFINITION Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-9.
ACCESSION AB097584
VERSION AB097584.1 GI:31442174
KEYWORDS .
SOURCE Macaca fuscata (Japanese macaque)
ORGANISM Macaca fuscata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS Yi,J.M., Takenaka,O. and Kim,H.S.
TITLE Molecular characterization and phylogenetic relationship of HERV-W family in the Macaca fuscata
JOURNAL Arch. Virol. 148 (8), 1613-1622 (2003)
PUBMED 12898334
REFERENCE 2 (bases 1 to 539)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University, Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .539
/organism="Macaca fuscata"
/mol_type="genomic DNA"
/isolate="JM-WE-9"
/db_xref="taxon:9542"
repeat_region 1. .539
/note="HERV-W env"
/rpt_family="HERV-W"

Query Match 6.2%; Score 467.1; DB 1; Length 539;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 498; Conservative 3; Mismatches 37; Indels 1; Gaps 1;

```
Qy       6200 ACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATS 6259
              |||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||:
Db          2 ACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGAACCTCTTATTTCCAATC 61

Qy       6260 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
              |||  ||||||||| |||||||||||| |||||||||||||||||| |||| | ||||||
Db         62 TGGGGATAACCCATGCCTCAAACCTCAGCTGTGTAAAATTTAGCAACACTATAGACACAA 121

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
              |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db        122 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCTCACACGAATAGTCTGCCTACCCTCAG 181

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAATAGTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTT 241

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
              ||||||||| |||||||||||||||||| |||:||||||||||||||| ||||||||||||
Db        242 TGTGCTTCCCCTCATTCTTAGTGCCCCC-ATGACCATCTACACTGAATAAGATTTATACA 300

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
              |||||| ||| ||||||||| |||||||||||||||||||||||||| |||||| ||||
Db        301 TTTATGTTATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTATGTTATTGGAG 360

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
              ||||||||||||   |||| |||||||| ||||| |||||||||||||||||||||||||
```

```
Db        361 AAGGAGTGCTAGGCAGACTAAGTACTGGCGTTGGCAGTATCACAACCTCTACTCAGTTCT 420

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
              |||| |||||||||||| |||||||| ||||||||| |||| || ||||| |||||||
Db        421 ACTATAAACTATCTCAAGGACTAAATGGTGACATGGAATGGGTTGCTGACTCACTGGTCA 480

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| ||||||||||||||||||||||||||||||:|||||||||||||||||
Db        481 CCTTGCAAGATTAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTA 539
```

RESULT 696
AB115114
LOCUS AB115114 533 bp RNA linear PRI 01-FEB-2006
DEFINITION Homo sapiens RNA transcript, HERV-W env, isolate:CWE5-3.
ACCESSION AB115114
VERSION AB115114.1 GI:41016549
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yi,J.M., Kim,H.M. and Kim,H.S.
TITLE Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells
JOURNAL J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
PUBMED 15105536
REFERENCE 2 (bases 1 to 533)
AUTHORS Yi,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2003) Joo-Mi Yi, Pusan National University, Biology; 30, Changjeon-dong, Keumjeong-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346, Fax:82-5158-12962)
FEATURES Location/Qualifiers
source 1. .533
/organism="Homo sapiens"
/mol_type="other RNA"
/isolate="CWE5-3"
/db_xref="taxon:9606"
/cell_line="TE-1"
/note="HERV-W env"
repeat_region 1. .533
/note="HERV-W env"

Query Match 6.2%; Score 465.5; DB 1; Length 533;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 492; Conservative 3; Mismatches 28; Indels 3; Gaps 1;

```
Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
              | | |||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db         11 ACACATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
              |||||||||||||||||||||||||||||||||||  | ||||||||||||||||||
Db         71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        131 CATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 190

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        191 CTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCCC 250

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
              |||||||||| ||||:||| ||||||||||||||||||||||||| | |||||||| |
Db        251 ATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACC 310

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
```

```
Db        311 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 370

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
              | ||||||||||||||  |||||||||||| | ||||||||||   ||||||||||||
Db        371 CGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT---CTACAAACTATC 427

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
              ||||||||| ||||| ||||||||| |||| ||||||||||||||||||||||||||||
Db        428 TCAAGAACTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCA 487

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||| ||||||| |||||||||||:|||| ||||||||||
Db        488 ACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAGCTTTA 533


RESULT 697
AB096029
LOCUS       AB096029                543 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-3.
ACCESSION   AB096029
VERSION     AB096029.1  GI:30698505
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-3"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .543
                     /gene="env"
     repeat_region   1. .543
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.2%;  Score 465.2;  DB 1;  Length 543;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches  500;  Conservative    3;  Mismatches   26;  Indels    2;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              |||| |||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 TTGGAGATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||||| || ||||| |||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTCAAGGGCTCATCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              || |||||||||||||||||||||||||:|||:|||||||||||||||| |||||||||
Db        241 ATATGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAGGATTTATAC 300

Qy       6499 AGTTATGTCATATC-TAAGCCCCGCAAC-AAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
              | | ||||| || | |||||||||||| ||||||||||||||||||||||||||||| |
Db        301 AATCATGTCGTACCTTAAGCCCCGCAACAAAAAGAGTACCCATTCTTCCTTTTGTTATCG 360

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
              ||||| |||||||||| | ||||||| |||||||||| ||||||||||||||||||||||
Db        361 GAGCAAGAGTGCTAGGCGGACTAGGTTCTGGCATTGGTGGTATCACAACCTCTACTCAGT 420

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
              |||||||||||||||||||||||||||||| |||| |||||||||||| ||||||||||
Db        421 CCTACTACAAACTATCTCAAGAACTAAATGGTGACACGGAACGGGTCGCTGACTCCCTGG 480

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATC 6727
              |||| ||||||||||||||||||||||||||||||||||||||:|||||
Db        481 TCACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATC 531


RESULT 698
AB055414
LOCUS       AB055414                536 bp    DNA     linear   PRI 07-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:OVCAR-1.
ACCESSION   AB055414
VERSION     AB055414.1  GI:15128496
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Lee,W.-H., Kim,H.-M. and Kim,H.-S.
  TITLE     Identification of new endogenous retroviral sequences belonging to
            the HERV-W family in human cancer cells
  JOURNAL   Intervirology (2001) In press
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="OVCAR-1"
                     /db_xref="taxon:9606"
                     /cell_line="NIH:OVCAR-3"
                     /note="HERV-W env"
     repeat_region   1. .536
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.2%;  Score 464.4;  DB 1;  Length 536;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   30;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||| ||| ||| ||||||||||||||||||||
Db        118 ACCAACTCCCAATGCATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCA 176

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
```

```
                ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        177 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCT 236

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||| |||||||||||||||||:|||:|||||||||||||||||||| ||||
Db        237 ATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                | | |||||||| |||| ||| ||||||||||||||||||||||||||||||||| |||
Db        297 AATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                |||||||||||||| | | |||||||||||| |||||||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||| ||||||| ||||||| |||||||||| |||||||||| |||||||||
Db        417 TACTACAAACTATTTCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||  |||||||||||||||| |||||||||||||||||||:||||| |||||| ||||
Db        477 ACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 536


RESULT 699
AB071980
LOCUS       AB071980                539 bp    DNA     linear   PRI 26-MAR-2002
DEFINITION  Macaca fuscata DNA, HERV-W envelope, isolate:WE-JP2.
ACCESSION   AB071980
VERSION     AB071980.1  GI:19716044
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Kim,H. and Jeon,S.
  TITLE     Isolation of the Endogenous Retrovirus W family in Japanese monkey
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="WE-JP2"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W envelope"

  Query Match             6.2%;  Score 463.9;  DB 1;  Length 539;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  496;  Conservative    3;  Mismatches   39;  Indels    1;  Gaps    1;

Qy       6200 ACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATS 6259
                ||||||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||:
Db          2 ACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGAACCTCTTATTTCCAATC 61

Qy       6260 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
                |||  ||||| ||| |||||||||||| |||||||||||||||||| |||| | ||||||
Db         62 TGGGGATAACTCATGCCTCAAACCTCAGCTGTGTAAAATTTAGCAACACTATAGACACAA 121

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
                |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db        122 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCTCACACGAATAGTCTGCCTACCCTCAG 181

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
                ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAATAGTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTT 241

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
                ||||||||| |||||||||||||||||| |||:|||||||||||||| ||||||||||||
Db        242 TGTGCTTCCCCTCATTCTTAGTGCCCCC-ATGACCATCTACACTGAATAAGATTTATACA 300

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
                ||||||  ||| ||||||||| |||||||||||||||||||||||||| ||||||  ||||
Db        301 TTTATGTTATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTATGTTATTGGAG 360

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
                |||||||||||||   |||| ||||||| |||||  |||||||||||||||||||||||||
Db        361 AAGGAGTGCTAGGCAGACTAAGTACTGGCGTTGGCAGTATCACAACCTCTACTCAGTTCT 420

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
                ||||  |||||||||||||| |||||||||  |||||||||  |||| || |||||  |||||||
Db        421 ACTATAAACTATCTCAAGGACTAAATGGTGACATGGAATGGGTTGCTGACTCACTGGTCA 480

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||||  |||||||||||||||||||||||  ||||||:|||||||||||||||
Db        481 CCTTGCAAGATTAACTTAACTCCCTAGCAGCAGTATTCCTTCAAAATCGAAGAGCTTTA 539


RESULT 700
AB097577
LOCUS       AB097577                539 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-2.
ACCESSION   AB097577
VERSION     AB097577.1  GI:31442167
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-2"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.2%;  Score 463.9;  DB 1;  Length 539;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  496;  Conservative    3;  Mismatches   39;  Indels    1;  Gaps    1;

Qy       6200 ACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATS 6259
                ||||||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||:
Db          2 ACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGAACCTCTTATTTCCAATC 61

Qy       6260 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
                |||  ||||| ||| |||||||||||| |||||||||||||||||| |||| | ||||||
Db         62 TGGGGATAACTCATGCCTCAAACCTCAGCTGTGTAAAATTTAGCAACACTATAGACACAA 121

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
                |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
```

```
Db        122 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCTCACACGAATAGTCTGCCTACCCTCAG 181

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAATAGTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTT 241

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
              ||||||||| |||||||||||||||||| |||:||||||||||||||| |||||||||||||
Db        242 TGTGCTTCCCCTCATTCTTAGTGCCCCC-ATGACCATCTACACTGAATAAGATTTATACA 300

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
              |||||| ||| ||||||||| |||||||||||||||||||||||||| |||||| ||||
Db        301 TTTATGTTATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTATGTTATTGGAG 360

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
              ||||||||||||   |||| |||||||| ||||| ||||||||||||||||||||||||
Db        361 AAGGAGTGCTAGGCAGACTAAGTACTGGCGTTGGCAGTATCACAACCTCTACTCAGTTCT 420

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
              |||| |||||||||||| |||||||||| |||| || ||||| |||||||
Db        421 ACTATAAACTATCTCAAGGACTAAATGGTGACATGGAATGGGTTGCTGACTCACTGGTCA 480

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||| |||||||||||||||||||| |||||||:||||||||||||||||
Db        481 CCTTGCAAGATTAACTTAACTCCCTAGCAGCAGTATTCCTTCAAAATCGAAGAGCTTTA 539


RESULT 701
AB114958
LOCUS       AB114958                535 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE6-3.
ACCESSION   AB114958
VERSION     AB114958.1  GI:41016466
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE6-3"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.2%;  Score 463.3;  DB 1;  Length 535;
  Best Local Similarity   93.2%;  Pred. No. 0;
  Matches  490;  Conservative    4;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
              | | |||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db         11 ACACATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70

Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
              |||||||||||||||||||||||||||||||||| | | ||||||||||||||||||
Db         71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
              |||||||||||||||||| ||| ||| ||||||||||||||||||||||||||||||||
Db         131 CATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 189

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         190 CTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTGTC 249

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
              |||||||||||||||:|||:|||||||||||||||||||||| ||||| | |||||||| |
Db         250 ATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATACAATCATGTCATACC 309

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
              ||| ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         310 TAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 369

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
              | | |||||||||||| |||||||||||||||||| ||||||||||||||||||||||||
Db         370 CGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTCTACTACAAACTATT 429

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
              ||||||| ||||||| |||||||||| |||||||||| |||||||||||  |||||||||
Db         430 TCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTCACCGAGCAAGATCA 489

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||| ||||||||||||||||||||:||||| ||||| ||||
Db         490 ACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 535


RESULT 702
AB114940
LOCUS       AB114940                535 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-16.
ACCESSION   AB114940
VERSION     AB114940.1  GI:41016448
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-16"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.2%;  Score 461.7;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  489;  Conservative    4;  Mismatches   32;  Indels    1;  Gaps    1;

Qy       6213 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATSTGGAAATAACCCA 6272
              | | |||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db         11 ACACATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATAACCCA 70
```

```
Qy       6273 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCAATG 6332
             ||||||||||||||||||||||||||||||||||| | | ||||||||||||||||||||
Db         71 TACCTCAAACCTCACCTGTGTAAAATTTAGCAATACGATAGACACAACCAACTCCCAATG 130

Qy       6333 CATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 6392
             |||||||||||||||||| ||| ||| ||||||||||||||||||||||||||||||||||
Db        131 CATCAGGTGGGTAACTCC-CCCCCACGAATAGTCTGCCTACCCTCAGGAATATTTTTTGT 189

Qy       6393 CTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTC 6452
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||| ||
Db        190 CTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTGTC 249

Qy       6453 ATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATC 6512
             ||||||||||||||:|||:|||||||||||||||||| ||||| | |||||||| |
Db        250 ATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTCATACAATCATGTCATACC 309

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             ||| ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        310 TAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGG 369

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
              | | | |||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db        370 CGGAATGGGTACTGGCACTGGCGGTATCACAACCTGTACTCAGTTCTACTACAAACTATT 429

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             ||||||| ||||||| |||||||||| ||||||||| |||||||||||  |||||||||
Db        430 TCAAGAAATAAATGGTGACATGGAACAGGTCGCCGATTCCCTGGTCACCGAGCAAGATCA 489

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||  ||||||||||||||||||||:||||| ||||| ||||
Db        490 ACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGTTTTA 535


RESULT 703
AB050997
LOCUS       AB050997               532 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-3-6.
ACCESSION   AB050997
VERSION     AB050997.1  GI:14020820
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 532)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-3-6"
                     /db_xref="taxon:9606"
                     /chromosome="3"
     repeat_region   1. .532
                     /rpt_family="HERV-W env"

  Query Match             6.2%;  Score 461.6;  DB 1;  Length 532;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   28;  Indels    8;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACATCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||     | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAAT-----AGACACA 115

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         116 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTGCCCTCA 175

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||| || |||||||||||||||||| ||||||||   |||||||||||||||||||||||||
Db         176 GGAGTACTTTTTGTCTGTGGTACCTTAGCCTATC---GTTTGAATGGCTCTTCAGAATCT 232

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||   |||||||: ||:|||||||||||||||||||||||||||||
Db         233 ATGTGCTTCCTCTCATTCTCCATGCCCCCTGTGACCATCTACACTGAACAAGATTTATAC 292

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | |||||||||| ||||||||| |||||||||||||||||||||||||||||||||| |||
Db         293 AATTATGTCATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGA 352

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||| |||||||||| | ||||||| |||||||||| |||||||||||||||||||||||| |
Db         353 GCAAGAGTGCTAGGCGGACTAGGTTCTGGCATTGGTGGTATCACAACCTCTACTCAGTCC 412

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||||||||||||||||||||||| ||||||||| ||| ||| ||||||| |||
Db         413 TACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACAGGTTGCCAACTCCCTCGTC 472

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||||||||||||||||||||||||||||||||||:||||| ||| ||||||
Db         473 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGCGCTTTA 532


RESULT 704
AB096033
LOCUS       AB096033               540 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-9.
ACCESSION   AB096033
VERSION     AB096033.1  GI:30698509
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-9"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .540
                     /gene="env"
     repeat_region   1. .540
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"
```

Query Match 6.2%; Score 461.2; DB 1; Length 540;
Best Local Similarity 92.8%; Pred. No. 0;
Matches 502; Conservative 2; Mismatches 35; Indels 2; Gaps 2;

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              |||| |||||||||||||||||||||||||||||||||||||||||||| ||| | |||||
Db         61 TTGGAGATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATTCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||| |||||||||||||||||| |||||||||||||||||||| ||||
Db        181 GGAATATTTTTTGTTTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGGATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              ||||| |||||||||||||||| |||||| |||: ||||||||||||||||||||||||
Db        241 ATGTGGTTCCTCTCATTCTTAGCGCCCCC-ATGATCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | |||||||||| ||||| ||| |||||||||||||||||||||| |||||||||| |||
Db        300 AATTATGTCATACCTAAGTCCCATAACAAAAGAGTACCCATTCTTTCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||  | ||||||| |||||||||| ||||||| ||||||||||||||||
Db        360 GCAGGAGTGCTAGACGGACTAGGTGCTGGCATTGGTGGTATCAAAACCTCTACTCAGTTC 419

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGA-ACGGGTCGCCGACTCCCTGGT 6677
              ||||| |||||||||||||||||||||| ||||||||   ||||| || |||| |||||
Db        420 TACTATAAACTATCTCAAGAACTAAATGGTGACATGGATTTGGGTCACCAACTCACTGGT 479

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              |||||||||||||||||||||||||||||||||||||||||| | :|||||||||||||||
Db        480 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTAAAATCGAAGAGCTTT 539

Qy       6738 A 6738
              |
Db        540 A 540
```

```
RESULT 705
CQ104355
LOCUS       CQ104355                568 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 13214 from Patent WO0157272.
ACCESSION   CQ104355
VERSION     CQ104355.1  GI:41073406
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 13214 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 2.7"
```

Query Match 6.2%; Score 460.7; DB 1; Length 568;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 494; Conservative 28; Mismatches 31; Indels 3; Gaps 3;

```
Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||   |  | |||||||||||| |||||||||||| |||||||||||||||||||||||||
Db         16 ATTCTAGGAGTACAGAAACCCAATGGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTAT 74

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              || | |||||||||||||| | |||||||||||||||||| |||||||||| ||::|||||
Db         75 CAGT-GAGGCCGTTGTCCCTCTATACCCAGCTGTACCTAACCCTTATACTCTGCTTTCCC 133

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              |||||||| |||||||||||||||||| | ||||||||||:|||||||||| |||||||
Db        134 AAATACCATAGGAAGCAGAGTGGTTTATAATCCTGGACCTTAAGGATGCCTTTTTCTGCA 193

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              |||||||||||||||||||| |||||| |||||||||||||| |||| ||||||:| |
Db        194 TCCCTGTACATCCTGACTCTCGATTCTTTTTTGCCTTTGAAGATCCTTCGAACCCAACGT 253

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCC 2934
              |||||||||| |||||||:|||||||||||||||||||||||||||||:|||||||||||||
Db        254 CTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCC 313

Qy       2935 AGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKG 2994
              || |||||||||||||||||:||:|::|||||||||||||||||||||||||:||| |
Db        314 AGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACACTCTTGTCCTTCAGTAAG 373

Qy       2995 TGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRC 3054
              || |||||||||||| :||:|||::|||||||||||||||||||||||||||||||||:|
Db        374 TGAATGATTTACTTTCAGCTGCCCGTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCAC 433

Qy       3055 TCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCT 3114
              ||||:|||||||| ||:|||||||||||:|||||||||:| |:|||||:|||||||
Db        434 TCTTAAATTTCCTTGCCACCTGTGGCTACAAGGTTTCCAAACCA-AAGCTCAACTCTGCT 492

Qy       3115 CACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGG 3174
              || | |||| ||||||||||||:||||:|||||||||||||||:|||||||||||||||
Db        493 CAGAACAGGCTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAGGGCCCTCAGTGAGG 552

Qy       3175 AAYRYATCCAGCCTAT 3190
              ||:::|||||||| ||
Db        553 AACGTATCCAGCCCAT 568
```

```
RESULT 706
AB079506
LOCUS       AB079506                536 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-6.
ACCESSION   AB079506
VERSION     AB079506.1  GI:22122080
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-6"
                     /db_xref="taxon:9606"
```

```
                     /tissue_type="brain"
     repeat_region   1. .536
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 459.6;  DB 1;  Length 536;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    4;  Mismatches   33;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             || |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||| |||||||||||||||||||||||||||||||||||||||||| |   |||
Db         61 CTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGCAACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||| |||| |||||||||| | |||||||||||||| |||||
Db        121 GCCAACTCCCAATGCATCAGGCGGGTTACTCCTCCCACGCGAATAGTCTGCCTAACCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:|||:|||||||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTG-CCCCTATGACCATCTACACTGAACAAAATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||   | |||||||||||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | |||||||||||||  |||||||||||||| | ||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 418

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||||||||||||||| ||||||||||||||||| ||||||||||||
Db        419 --CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              || |||||||||||||||||||||||||||||| ||||||||||:|||| ||||||||||
Db        477 ACATTGCAAGATCAACTTAACTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 707
AB096036
LOCUS       AB096036                536 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-13.
ACCESSION   AB096036
VERSION     AB096036.1  GI:30698512
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-13"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .536
                     /gene="env"
     repeat_region   1. .536
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 459.6;  DB 1;  Length 536;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    4;  Mismatches   33;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             || |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||| |||||||||||||||||||||||||||||||||||||||||| |   |||
Db         61 CTGGAAATAATCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGCAACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||| |||| |||||||||| | |||||||||||||| |||||
Db        121 GCCAACTCCCAATGCATCAGGCGGGTTACTCCTCCCACGCGAATAGTCTGCCTAACCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||| ||||:|||:|||||||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTG-CCCCTATGACCATCTACACTGAACAAAATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||   | |||||||||||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | |||||||||||||  |||||||||||||| | ||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 418

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                |||||||||||||||||||||||||| ||||||||||||||||| ||||||||||||
Db        419 --CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              || |||||||||||||||||||||||||||||| ||||||||||:|||| ||||||||||
Db        477 ACATTGCAAGATCAACTTAACTCCCTAGCAGCGGTAGTCCTTCAAAATTGAAGAGCTTTA 536


RESULT 708
AB096030
LOCUS       AB096030                539 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-4.
ACCESSION   AB096030
VERSION     AB096030.1  GI:30698506
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
```

```
          Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-4"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .539
                     /gene="env"
     repeat_region   1. .539
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 458.5;  DB 1;  Length 539;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches  506;  Conservative    3;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||| ||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACCCCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAA-CCTCACCTGTGTAAAA-TTTAGCAATACTACATACA 6316
              :|||||||||||||  ||||||| |||||||||||||||| |||||||||||||| | |||
Db         58 CTGGAAATAACCCAATCCTCAAACCCTCACCTGTGTAAAATTTTAGCAATACTATAGACA 117

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        118 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCT 177

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        178 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAAT 237

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
              || |||||||||||||||||||||| ||||:||| ||||||||||||||||||||||||
Db        238 CTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTAT 297

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db        298 ACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTG 357

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             |||||||||||||||| | | ||| |||||||||||||||||||| ||||||||||||||||
Db        358 GAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCGCAACCTCTACTCAGT 417

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             |||||||||||||| |||||||||||||||| ||||||||| ||||||| ||   ||||||
Db        418 TCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTGG 477

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             |||||||||||||||||||||||||||||||||||||||||||:|||||||||| ||||
Db        478 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGCGCTT 537

Qy       6737 TA 6738
              ||
Db        538 TA 539


RESULT 709
AB071979
LOCUS       AB071979                543 bp    DNA     linear   PRI 26-MAR-2002
DEFINITION  Macaca fuscata DNA, HERV-W envelope, isolate:WE-JP1.
ACCESSION   AB071979
VERSION     AB071979.1  GI:19716043
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Kim,H. and Jeon,S.
  TITLE     Isolation of the Endogenous Retrovirus W family in Japanese monkey
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="WE-JP1"
                     /db_xref="taxon:9542"
     repeat_region   1. .543
                     /note="HERV-W envelope"

  Query Match             6.1%;  Score 458.5;  DB 1;  Length 543;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACGGAAATAAACTCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||| |||||||||||||||||||| |||||||| |||||||||||||| |||||
Db         61 GTGGAAATAGCCCATACCTCAAACCTCACCTATGTAAAATGTAGCAATACTACAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||  ||||| ||||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTTGCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||| |||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGTCTATGGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || |||||||||||||||| ||||||| |||||||||||||||||||| |||
Db        300 AATCATGTCTTACCTAAGCCCCGCAACAGAAGAGTATCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGG---TACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              |||||||||||||| | ||||||   ||||||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTATTACTGGCATTGGCGGTATCACAACCTCTACTCAG 419

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              ||||||||| |||||||||||||| |||||||| ||||||||| |||| ||||||||||||
Db        420 TTCTACTACGAACTATCTCAAGAGCTAAATGGTGACATGGAATGGGTTGCCGACTCCCTG 479

Qy       6676 GTCACCTTGCAAGATC-AACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
              |||||||||||||||| |||||||||||||| ||| ||||||||||:||||| ||||||
Db        480 GTCACCTTGCAAGATCAAACTTAACTCCCTAGAAGCTGTAGTCCTTCGAAATCAAAGAGC 539

Qy       6735 TTTA 6738
              ||||
Db        540 TTTA 543


RESULT 710
AB097576
LOCUS       AB097576                543 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-1.
ACCESSION   AB097576
VERSION     AB097576.1  GI:31442166
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE  1
  AUTHORS  Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE    Molecular characterization and phylogenetic relationship of HERV-W
           family in the Macaca fuscata
  JOURNAL  Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED  12898334
REFERENCE  2  (bases 1 to 543)
  AUTHORS  Yi,J. and Kim,H.
  TITLE    Direct Submission
  JOURNAL  Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
           Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
           Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
           Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-1"
                     /db_xref="taxon:9542"
     repeat_region   1. .543
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 458.5;  DB 1;  Length 543;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  506;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACGGAAATAAACTCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||| |||||||||||||||||||| |||||||| |||||||||||||| |||||
Db         61 GTGGAAATAGCCCATACCTCAAACCTCACCTATGTAAAATGTAGCAATACTACAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||  ||||| ||||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTTGCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||| |||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGTCTATGGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| || |||||||||||||| ||||||| |||||||||||||||||||| |||
Db        300 AATCATGTCTTACCTAAGCCCCGCAACAGAAGAGTATCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGG---TACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              |||||||||||||| | ||||||   |||||||||||||||||||||||||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTATTACTGGCATTGGCGGTATCACAACCTCTACTCAG 419

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              ||||||||| |||||||||||||| |||||||| ||||||||| |||| ||||||||||||
Db        420 TTCTACTACGAACTATCTCAAGAGCTAAATGGTGACATGGAATGGGTTGCCGACTCCCTG 479

Qy       6676 GTCACCTTGCAAGATC-AACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
              |||||||||||||||| |||||||||||||| ||| ||||||||||:||||| ||||||
Db        480 GTCACCTTGCAAGATCAAACTTAACTCCCTAGAAGCTGTAGTCCTTCGAAATCAAAGAGC 539

Qy       6735 TTTA 6738
              ||||
Db        540 TTTA 543


RESULT 711
AB071981
LOCUS       AB071981                 535 bp    DNA     linear   PRI 26-MAR-2002
DEFINITION  Macaca fuscata DNA, HERV-W envelope, isolate:WE-JP3.
ACCESSION   AB071981
VERSION     AB071981.1  GI:19716045
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Kim,H. and Jeon,S.
  TITLE     Isolation of the Endogenous Retrovirus W family in Japanese monkey
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="WE-JP3"
                     /db_xref="taxon:9542"
     repeat_region   1. .535
                     /note="HERV-W envelope"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||||||||||||||||||||| |||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||||||||||   |||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||| |||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 712
AB097578
LOCUS       AB097578                 535 bp    DNA     linear   PRI 05-AUG-2003
```

```
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-3.
ACCESSION   AB097578
VERSION     AB097578.1  GI:31442168
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-3"
                     /db_xref="taxon:9542"
     repeat_region   1. .535
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  || ||||||||||||||| |||||||||||||||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535
```

```
RESULT 713
AB126133
LOCUS       AB126133                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE1-3.
ACCESSION   AB126133
VERSION     AB126133.1  GI:47169733
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE1-3"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  || ||||||||||||||| |||||||||||||||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||||
```

```
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 714
AB126136
LOCUS       AB126136                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE2-1.
ACCESSION   AB126136
VERSION     AB126136.1  GI:47169736
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE2-1"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| |||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || ||||||||||||| |||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476


Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 715
AB126141
LOCUS       AB126141                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE3-1.
ACCESSION   AB126141
VERSION     AB126141.1  GI:47169741
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE3-1"
                     /db_xref="taxon:9542"
                     /tissue_type="kidney"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| |||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || ||||||||||||| |||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCGCTACTCAGTTC 416
```

```
Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 716
AB126145
LOCUS       AB126145                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE5-1.
ACCESSION   AB126145
VERSION     AB126145.1  GI:47169745
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE5-1"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 717
AB126148
LOCUS       AB126148                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE6-2.
ACCESSION   AB126148
VERSION     AB126148.1  GI:47169748
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE6-2"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 456.1;  DB 1;  Length 535;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  503;  Conservative    4;  Mismatches   28;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||| |||||||||||||||||| |||
```

```
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 718
AB071982
LOCUS       AB071982                539 bp    DNA     linear   PRI 26-MAR-2002
DEFINITION  Macaca fuscata DNA, HERV-W envelope, isolate:WE-JP5.
ACCESSION   AB071982
VERSION     AB071982.1  GI:19716046
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Kim,H. and Jeon,S.
  TITLE     Isolation of the Endogenous Retrovirus W family in Japanese monkey
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="WE-JP5"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W envelope"

  Query Match             6.1%;  Score 455.7;  DB 1;  Length 539;
  Best Local Similarity   93.2%;  Pred. No. 0;
  Matches  504;  Conservative    3;  Mismatches   31;  Indels    3;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||| |||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAGACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||| ||| |||||||||| ||||||||| | |||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGATAAATCCTCCCACACGAATAGTCTGACCACCCTCA 180

Qy       6379 GGAATA-TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
              | |||| |||||||||||||||| |||||| ||| ||||||||||||||||||||||||
Db        181 GAAATATTTTTTTGTCTGTGGTATCTCAGCTTATTGTTGTTTGAATGGCTCTTCAGAATC 240

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
              |||||||||||||||||| |||||||||| |||:|||| ||||||||||||||||||||
Db        241 TATGTGCTTCCTCTCATTTTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATA 299

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
              ||  ||||||||||||||  | ||||| |||||||||||||||||||||||||||| ||

Db        300 CAACTATGTCATATCTAAGGGCTGCAAC-AAAGAGTACCCATTCTTCCTTTTGTTATTGG 358

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
              | | |||||||||||| |||||||||||||||||||||||||||||||||| ||||||
Db        359 AACGGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTAGTCAGTT 418

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
              |||||||||||||||||||||||||||||||| |||| ||||| |||||| ||||||||||
Db        419 CTACTACAAACTATCTCAAGAACTAAATGGTGACAGGGAACTGGTCGCTGACTCCCTGGT 478

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        479 TACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 538

Qy       6738 A 6738
              |
Db        539 A 539


RESULT 719
AB097579
LOCUS       AB097579                539 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-4.
ACCESSION   AB097579
VERSION     AB097579.1  GI:31442169
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-4"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 455.7;  DB 1;  Length 539;
  Best Local Similarity   93.2%;  Pred. No. 0;
  Matches  504;  Conservative    3;  Mismatches   31;  Indels    3;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||| |||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAGACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||| ||| |||||||||| ||||||||| | |||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGATAAATCCTCCCACACGAATAGTCTGACCACCCTCA 180

Qy       6379 GGAATA-TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
              | |||| |||||||||||||||| |||||| ||| ||||||||||||||||||||||||
Db        181 GAAATATTTTTTTGTCTGTGGTATCTCAGCTTATTGTTGTTTGAATGGCTCTTCAGAATC 240
```

```
Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             |||||||||||||||||| |||||||||| |||:|||| ||||||||||||||||||||||
Db        241 TATGTGCTTCCTCTCATTTTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATA 299

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             ||  ||||||||||||||  | ||||| |||||||||||||||||||||||||||||| ||
Db        300 CAACTATGTCATATCTAAGGGCTGCAAC-AAAGAGTACCCATTCTTCCTTTTGTTATTGG 358

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
             | | ||||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||
Db        359 AACGGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTAGTCAGTT 418

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
             |||||||||||||||||||||||||||||| |||| ||||| |||||| ||||||||||
Db        419 CTACTACAAACTATCTCAAGAACTAAATGGTGACAGGGAACTGGTCGCTGACTCCCTGGT 478

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        479 TACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 538

Qy       6738 A 6738
             |
Db        539 A 539
```

```
RESULT 720
AB126155
LOCUS       AB126155                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-5.
ACCESSION   AB126155
VERSION     AB126155.1  GI:47169755
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-5"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 455.1;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||||||||||| |||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAACAATACTATAGACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||||  |||||||||||||||||| ||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535
```

```
RESULT 721
AB126163
LOCUS       AB126163                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE9-1.
ACCESSION   AB126163
VERSION     AB126163.1  GI:47169763
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE9-1"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 455.1;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
```

```
              :|||||||||||||||||||||| |||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||||||||||||||||||| |||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| |||||: ||:|||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| |  ||||||||| |||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||| |||||||||| |||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 722
AB114936
LOCUS       AB114936                536 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE2-12.
ACCESSION   AB114936
VERSION     AB114936.1  GI:41016444
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE2-12"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     repeat_region   1. .536
                     /note="HERV-W env"

  Query Match             6.1%;  Score 454.8;  DB 1;  Length 536;
  Best Local Similarity   91.9%;  Pred. No. 0;
  Matches  496;  Conservative    4;  Mismatches   36;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              || |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AATAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||| ||| ||||||||||||||||||||||||||||||||||| |   |||
Db         61 CTGGAAATAATCCACACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGCAACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||| |||| |||||||||| | ||||||||||||| |||||
Db        121 GCCAACTCCCAATGCACCAGGCGGGTTACTCCTCCCACGCGAATAGTCTGCCTAACCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| ||||:|||:|||||||||||||||| ||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTG-CCCCTATGACCATCTACACTGAACAAAATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||   | |||||||||||||||||||||||||||||||||||||||||| |||
Db        300 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 359

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              |||||||||||||| | ||||||||||||  |||||||||||||| | ||||||||||
Db        360 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTT- 418

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||||||||||||||||||| |||||||||||||||| ||||||||||
Db        419 --CTACAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              || |||||||||||||||||||||||||||||||||||||||||:|||||  | |||||||
Db        477 ACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTTTA 536


RESULT 723
AB054088
LOCUS       AB054088                533 bp    DNA     linear   PRI 06-DEC-2001
DEFINITION  Homo sapiens DNA, human endogenous retrovirus HERV-W family
            sequence, isolate:MCF7-1.
ACCESSION   AB054088
VERSION     AB054088.1  GI:15005018
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
  TITLE     Molecular cloning and phylogenetic analysis of new human endogenous
            retrovirus HERV-W family in cancer cells
  JOURNAL   Curr. Microbiol. (2001) In press
REFERENCE   2  (bases 1 to 533)
  AUTHORS   Lee,J. and Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .533
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="MCF7-1"
                     /db_xref="taxon:9606"
                     /cell_line="MCF7"
                     /note="HERV-W env"
     repeat_region   1. .533
                     /rpt_family="human endogenous retrovirus HERV-W"

  Query Match             6.1%;  Score 454.7;  DB 1;  Length 533;
  Best Local Similarity   93.0%;  Pred. No. 0;
```

```
  Matches  502;  Conservative    4;  Mismatches   27;  Indels    7;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||| ||||   ||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAAT 57

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         58 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 117

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        118 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 177

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        178 GGAATATTTTTTGTCTGTGGTACTTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||:|||:|||||||||||||||||| ||||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGTG-CCCCTATGACCATCTACACTGAACAAAATTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||   | ||||||||||||||||||||||||||||||||||||||||||| |||
Db        297 AATCATGTTGCACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||| | ||||||||||||  |||||||||||||| | ||||||||||||
Db        357 GCAGGAGTGCTAGGCGGACTAGGTACTGGCCCTGGCGGTATCACAGCATCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||   |||||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        417 TA---CAAACTATCTCAAGAACTAAATGGTGACATGGAACGGGTCGCTGACTCCCTGGTC 473

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             || |||||||||||||||||||||||||||||||||||||||:|||||  | |||||||
Db        474 ACATTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAGACAGCTTTA 533


RESULT 724
AB097582
LOCUS       AB097582                 535 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-7.
ACCESSION   AB097582
VERSION     AB097582.1  GI:31442172
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-7"
                     /db_xref="taxon:9542"
     repeat_region   1. .535
                     /note="HERV-W env"
                     /rpt_family="HERV-W"
```

```
  Query Match             6.1%;  Score 454.5;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||| |||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAGAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||||||||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             |||||||||||||||   || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| ||||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:|||||  ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 725
AB126139
LOCUS       AB126139                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE2-6.
ACCESSION   AB126139
VERSION     AB126139.1  GI:47169739
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE2-6"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
```

```
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match              6.1%;  Score 454.5;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||| ||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAGCCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 726
AB126154
LOCUS       AB126154                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-4.
ACCESSION   AB126154
VERSION     AB126154.1  GI:47169754
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-4"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match              6.1%;  Score 454.5;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||| ||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAGCCTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||| |||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 727
AB126164
LOCUS       AB126164                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE9-3.
ACCESSION   AB126164
VERSION     AB126164.1  GI:47169764
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
```

```
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE9-3"
                     /db_xref="taxon:9542"
                     /tissue_type="ovary"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 454.5;  DB 1;  Length 535;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   29;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||| |||| ||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||  |||||||||||||||||||||||||||||  ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             ||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             ||||||||||||||||||||||| ||||||: ||:||||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||||  |||||||||||||  |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||  ||||||||||||  ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||||  |||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||||||  ||||||||||  |||||||||||||||||||:|||||  ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 728
AB097580
LOCUS       AB097580                539 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-5.
ACCESSION   AB097580
VERSION     AB097580.1  GI:31442170
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-5"
                     /db_xref="taxon:9542"
     repeat_region   1. .539
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.1%;  Score 454.1;  DB 1;  Length 539;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  503;  Conservative    3;  Mismatches   32;  Indels    3;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCTAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCTAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAGACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||| |||||||||||  ||| |||||||||| ||||||| | |||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGATAAATCCTCCCACACGAATAGTCTGACCACCCTCA 180

Qy       6379 GGAATA-TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
             | |||| |||||||||||||||| |||||| ||| ||||||||||||||||||||||||||
Db        181 GAAATATTTTTTTGTCTGTGGTATCTCAGCTTATTGTTGTTTGAATGGCTCTTCAGAATC 240

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             |||||||||||||||||| |||||||||| |||:|||| ||||||||||||||||||||||
Db        241 TATGTGCTTCCTCTCATTTTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATA 299

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             ||  |||||||||||||||  | ||||| ||||||||||||||||||||||||||||| ||
Db        300 CAACTATGTCATATCTAAGGGCTGCAAC-AAAGAGTACCCATTCTTCCTTTTGTTATTGG 358

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
             | | |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        359 AACGGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTAGTCAGTT 418

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
             |||||||||||||||||||||||||||||||||| |||| ||||| |||||| ||||||||||
Db        419 CTACTACAAACTATCTCAAGAACTAAATGGTGACAGGGAACTGGTCGCTGACTCCCTGGT 478

Qy       6678 CACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        479 TACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTT 538

Qy       6738 A 6738
              |
Db        539 A 539


RESULT 729
CQ101825/c
LOCUS       CQ101825                548 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10684 from Patent WO0157272.
ACCESSION   CQ101825
VERSION     CQ101825.1  GI:41070851
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
```

```
          analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 10684 09-AUG-2001;
          Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC021774.3
                     EXPRESSED IN PLACENTA, SIGNAL = 1.4"

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||: |||| ||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
              ||||:||||||||:|||||||||||||:||:|||::|||||| || |||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||:|||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
              ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 730
CQ140811/c
LOCUS       CQ140811                548 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10833 from Patent WO0157276.
ACCESSION   CQ140811
VERSION     CQ140811.1  GI:41098183
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 10833 09-AUG-2001;
          Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC021774.3
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.3"

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
              ||||||||:|||:|| :||||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
              :|||||: |||| ||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
              :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
              ||||:||||||||:|||||||||||||:||:|||::|||||| || |||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
              ||:||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
              ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
              ||::|||||||||||||||||||||||||||||||:|||||||| |||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
              |||||:|||||::||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
              |||||||:|||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
              ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 731
CQ224106/c
LOCUS       CQ224106                548 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10945 from Patent WO0157273.
ACCESSION   CQ224106
VERSION     CQ224106.1  GI:41206514
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
```

(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
JOURNAL Patent: WO 0157273-A 10945 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .548
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC021774.3
EXPRESSED IN ADULT LIVER, SIGNAL = 1.2"

Query Match 6.1%; Score 454; DB 1; Length 548;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 474; Conservative 47; Mismatches 27; Indels 4; Gaps 3;

```
Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :|||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| |||||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || |||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||:|||||||| |||||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||||: |||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1
```

RESULT 732
CQ262103/c
LOCUS CQ262103 548 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 10364 from Patent WO0157277.
ACCESSION CQ262103
VERSION CQ262103.1 GI:41234583
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 10364 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .548
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC021774.3
EXPRESSED IN FETAL LIVER, SIGNAL = 0.99"

Query Match 6.1%; Score 454; DB 1; Length 548;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 474; Conservative 47; Mismatches 27; Indels 4; Gaps 3;

```
Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :|||||||||||:|||||||| |||| |:  | :||||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| |||||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || |||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||||::||| ::||||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||:|||||||| |||||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||||: |||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1
```

RESULT 733
CQ299248/c
LOCUS CQ299248 548 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 10353 from Patent WO0186003.
ACCESSION CQ299248
VERSION CQ299248.1 GI:41259825
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung

```
  JOURNAL   Patent: WO 0186003-A 10353 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC021774.3
                     EXPRESSED IN LUNG, SIGNAL = 1.2"

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :|||||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| |||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||::||| ::|||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 734
CQ336270/c
LOCUS       CQ336270                548 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 10364 from Patent WO0157275.
ACCESSION   CQ336270
VERSION     CQ336270.1  GI:41285341
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 10364 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC021774.3
                     EXPRESSED IN BRAIN, SIGNAL = 1"

  Query Match             6.1%;  Score 454;  DB 1;  Length 548;
  Best Local Similarity   85.9%;  Pred. No. 0;
  Matches  474;  Conservative   47;  Mismatches   27;  Indels    4;  Gaps    3;

Qy       3235 TCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGG 3294
             ||||||||:|||:|| :|||||||||||:|||||||| |||| |:  | :||||||||||
Db        548 TCCTTGGCATAACAGCCTTCTGCCGAATATGGATTC-CCAGATACAGTGAAATAGCCAGG 490

Qy       3295 YCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYA 3354
             :|||||: |||| |||||||||||||||||||||||||||||||||| ||:||||||||:|
Db        489 CCATTATGTACATTAATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACA 430

Qy       3355 MCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGT 3414
             :||||| :::|||||||||||||||||||  ||  | | |||  ||||||| :||||||
Db        429 CCTGAAACAGAAGTGGCTTTCCAGGCCCTAAAGAAGGCCCT--AACCCAAACCCCAGTTG 372

Qy       3415 TAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCT 3474
             ||||:||||||||:|||||||||||||:||:|||::|||||| || ||||||:||:||||
Db        371 TAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAG-AAGAAACAGGAATAGCT 313

Qy       3475 CTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAG 3534
             ||:||||||||||| |||:|| :|||||:||||||||||::||| ::|||||:||||
Db        312 CTAGGAGTCCTTACGCAGGTCTGAGGGATGAGCTTGCAACCCGTGGTATACCTGAGTAAG 253

Qy       3535 GAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCA 3594
             ||||:|||||||||||||||||||:|:||||||||||:||||| |||| |||||||||
Db        252 GAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTAGCAGTAGCA 193

Qy       3595 GTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCT 3654
             ||::|||||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        192 GTCTTAGTATCTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTTCTGTGTGGACATCT 133

Qy       3655 CATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTA 3714
             |||||:|||||::|||||||||||||||||||||||||||||||||||||||:|||||
Db        132 CATGATGTGAACGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTA 73

Qy       3715 CTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACT 3774
             |||||||:|||||||||||||||||: ||||||||||:|||| ||||||||||||||
Db          72 CTTAAATATCAGGCTCTATTACTTGAAGAGCCAGTGCTGCGACTGCGCACTTGTGCAACT 13

Qy       3775 CTTAACCCAGYC 3786
             ||||||||||:|
Db          12 CTTAACCCAGCC 1


RESULT 735
AB126157
LOCUS       AB126157                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-1.
ACCESSION   AB126157
VERSION     AB126157.1  GI:47169757
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-1"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.1%;  Score 453.5;  DB 1;  Length 535;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   30;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||| |||||||||||||||||| |||| ||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:|||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 736
AB126150
LOCUS       AB126150                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE6-9.
ACCESSION   AB126150
VERSION     AB126150.1  GI:47169750
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE6-9"
                     /db_xref="taxon:9542"
                     /tissue_type="pancreas"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 452.9;  DB 1;  Length 535;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   30;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||| ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:|||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 737
AB126162
LOCUS       AB126162                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE8-9.
ACCESSION   AB126162
VERSION     AB126162.1  GI:47169762
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
```

```
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE8-9"
                     /db_xref="taxon:9542"
                     /tissue_type="stomach"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 452.9;  DB 1;  Length 535;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  501;  Conservative    4;  Mismatches   30;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||||||||||||  ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:|||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 738
AB120060
LOCUS       AB120060                526 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG10-7.
ACCESSION   AB120060
VERSION     AB120060.1  GI:45330731
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.

REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 526)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG10-7"
                     /isolation_source="spleen of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .526
                     /note="corresponds to gag"

  Query Match             6.0%;  Score 452.3;  DB 1;  Length 526;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches  493;  Conservative    9;  Mismatches    6;  Indels    5;  Gaps    4;

Qy        644 GACGTATTCTGGARAATT-GGGAMCAATTTGACCCTCAGACACTAAGAAAGAAA--CGAC 700
              |||||||||||||:|||| |||||:|||||||||||||| || |||||||||  |  ||||
Db         15 GACGTATTCTGGAGAATTGGGGACCAATTTGACCCTCAGGCAGTAAGAAAGAGAGCCGAC 74

Qy        701 TTATATTCTTCTGCA-GTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCAT 759
              |||||||||| |||| | ||||||||||||||| |||||||||||||||||||||||||
Db         75 TTATATTCTTGTGCAGGGGCCGCCTGGCACTCGTGAGGGAAGTATAAATTATAACACCAT 134

Qy        760 CTTACAGCTAGACYTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAAC 819
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CTTACAGCTAGACCTCTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAAC 194

Qy        820 TTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGG 879
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TTTCTTTTCATTAAGAGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGG 254

Qy        880 AAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGG 939
              |||||||||||||||||||||||||||||||||||||||||||||||||||::|||||||
Db        255 AAGCCTTCAGAGTCTACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGG 314

Qy        940 ACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAA 999
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ACCCCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAA 374

Qy       1000 AGAGTGCCAATATTCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGC 1059
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        375 AGAGTGCCAATATTCCCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGC 433

Qy       1060 CCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTA 1119
              |||||||||||||||||:||||::|||||||||||||||||||||||||||||||||||||
Db        434 CCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTA 493

Qy       1120 GGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              |||||||||||||||:|||||||||||||||
Db        494 GGTAAATTCTCAGATAACCCTGATGGCTATATT 526


RESULT 739
AB097581
LOCUS       AB097581                524 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-6.
ACCESSION   AB097581
VERSION     AB097581.1  GI:31442171
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 524)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .524
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-6"
                     /db_xref="taxon:9542"
     repeat_region   1. .524
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.0%;  Score 451.6;  DB 1;  Length 524;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  487;  Conservative    4;  Mismatches   33;  Indels   16;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||| || |||||||||||| ||||| | |||||||||||||||||||  | | |||
Db         61 CTGGAAACAATCCATACCTCAAATCTCACTTATGTAAAATTTAGCAATACTGTAGATACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTATTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             | |||||||||||| ||||||| ||||||:|||:||||||||||||||||||||||||||
Db        241 ACGTGCTTCCTCTCGTTCTTAGGGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||  | ||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCGAACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||| || ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGGGTACTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||| |||||||||  |||||  |||||  | |||
Db        421 TACTACAAAGTATCTCAAGAACTAAATGGTGACATGGAACAGGTCGCTGACTCTATAGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||                  |||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTG----------------CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 524


RESULT 740
AB097586
LOCUS       AB097586                524 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-11.
ACCESSION   AB097586
VERSION     AB097586.1  GI:31442176
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 524)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .524
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-11"
                     /db_xref="taxon:9542"
     repeat_region   1. .524
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             6.0%;  Score 451.6;  DB 1;  Length 524;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  487;  Conservative    4;  Mismatches   33;  Indels   16;  Gaps    1;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||| || |||||||||||| ||||| | |||||||||||||||||||  | | |||
Db         61 CTGGAAACAATCCATACCTCAAATCTCACTTATGTAAAATTTAGCAATACTGTAGATACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTATTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             | |||||||||||| ||||||| ||||||:|||:||||||||||||||||||||||||||
Db        241 ACGTGCTTCCTCTCGTTCTTAGGGCCCCCTATGACCATCTACACTGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | |||||  | ||||||||||||||||||||||||||||||||||||||||  ||
Db        301 AATCATGTCAAACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATCAGA 360

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||| || ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGGGGTACTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 420

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||| |||||||||  |||||  |||||  | |||
Db        421 TACTACAAAGTATCTCAAGAACTAAATGGTGACATGGAACAGGTCGCTGACTCTATAGTC 480

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             ||||||                  |||||||||||||||||||:||||| ||||||||||
Db        481 ACCTTG----------------CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 524


RESULT 741
AB126138
LOCUS       AB126138                535 bp    RNA     linear   PRI 14-MAY-2004
```

```
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE2-5.
ACCESSION   AB126138
VERSION     AB126138.1  GI:47169738
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE2-5"
                     /db_xref="taxon:9542"
                     /tissue_type="prostate"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 451.3;  DB 1;  Length 535;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   31;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||| ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||| ||||||: ||:||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| ||||||||||||||||||  ||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCAGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||   || |||||||||||| |||||||||||| ||||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||||||| ||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535
```

```
RESULT 742
AB096032
LOCUS       AB096032                544 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-7.
ACCESSION   AB096032
VERSION     AB096032.1  GI:30698508
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 544)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-7"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .544
                     /gene="env"
     repeat_region   1. .544
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             6.0%;  Score 451.2;  DB 1;  Length 544;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches  496;  Conservative    3;  Mismatches   41;  Indels    4;  Gaps    2;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              ||||||  || ||||||||||||| ||||| | |||||||||||||||||||  | | |||
Db         61 TTGGAAACAATCCATACCTCAAATCTCACTTATGTAAAATTTAGCAATACTGTAGATACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTTTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 240

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||||||||||:|||:|||||||||| ||||||||||||||||
Db        241 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACAGAACAAGATTTATAC 300

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAAC--AAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             | | ||||| || |||||||||    |  |||||||||||||||||||||||||||||
Db        301 AATCATGTCGTACCTAAGCCCCCACAACAAAAAGAGTACCCATTCTTCCTTTTGTTATCA 360

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             |||||||||||||||   |||||||||||||||||||| ||||||||||||||||||||||
Db        361 GAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGT 420

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACT--CCCT 6674
             |||||||||||||||||||| ||||||| |||||||||| |||| | ||||  ||
Db        421 TCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTTCCCTG 480
```

```
Qy       6675 GGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGC 6734
             |    ||||||||||||||||||||||||||||||||||||||||||:||||||| | |
Db        481 GTCACCCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAAAAC 540

Qy       6735 TTTA 6738
             ||||
Db        541 TTTA 544


RESULT 743
AB126144
LOCUS       AB126144                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE4-2.
ACCESSION   AB126144
VERSION     AB126144.1  GI:47169744
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE4-2"
                     /db_xref="taxon:9542"
                     /tissue_type="cerebellum"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 450.7;  DB 1;  Length 535;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  500;  Conservative    3;  Mismatches   32;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||| |||||||||||||||||| |||| ||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACNTCACCTGTGTAAAATTTAACAATGCTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||||: || ||||| |||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGCCCATCCACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             |||||||| |||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| |||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 744
AB126134
LOCUS       AB126134                 537 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE1-4.
ACCESSION   AB126134
VERSION     AB126134.1  GI:47169734
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE1-4"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.0%;  Score 450.7;  DB 1;  Length 537;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   32;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||| ||| |||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  ||||||||| |||||||||||||| |||||||||||||||||||| |||
```

```
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  || ||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||||||| |||||||||||||||| |||||||||| ||||||||||||||||||
Db        417 TACTACAAACTACCTCAAGAACTAAATGGTGACATGGAACAAGTCGCCGACTCCCTGGTC 476

Qy       6679 -ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTT 6737
               |||||||||||||||||||| ||||||||| |||||||||:|||| |||||||||
Db        477 AACCTTGCAAGATCAACTTAACTTCCTAGCAGCAATAGTCCTTCAAAATTGAAGAGCTTC 536

Qy       6738 A 6738
              |
Db        537 A 537


RESULT 745
AB096021
LOCUS       AB096021                543 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:FB-1.
ACCESSION   AB096021
VERSION     AB096021.1  GI:30698497
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hong,K., Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of HERV-W env Fragments in Fetal Brain cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Kim,H., Hong,K., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="FB-1"
                     /db_xref="taxon:9606"
                     /tissue_type="fetal brain"
     gene            1. .543
                     /gene="env"
     repeat_region   1. .543
                     /note="human endogenous retrovirus envelope protein"
                     /rpt_family="HERV-W"

  Query Match             6.0%;  Score 450.7;  DB 1;  Length 543;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  503;  Conservative    2;  Mismatches   35;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAA--TAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCA 6256
              ||||||||||||||||||   || ||||||||| |||| |||||| |||||||   | |
Db          1 AACAACTTCAGCACAGAAAATAAACCACCACTTCTGTTTAAGTAGGTCCTCTTGC--CAA 58

Qy       6257 ATSTGGAAATAACCCATACCTCAAACC-TCACCTGTGTAAAATTTAGCAATACTACATAC 6315
              | |||||||||||||||||||| |||||||||||||||||||||||||| | ||
Db         59 TCTGGAAAATAACCCATACCTCAAACCTTCACCTGTGTAAAATTTAGCAATACTATAGAC 118

Qy       6316 ACAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCC 6375
              ||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        119 ACAGCCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCC 178

Qy       6376 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
```

```
Db        179 TCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAA 238

Qy       6436 TCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTA 6495
              ||| |||||||||||||||||||||||| ||||:||| ||||||||||||||||||||||
Db        239 TCTGTGTGCTTCCTCTCATTCTTAGTGGCCCCTATGCCCATCTACACTGAACAAGATTTA 298

Qy       6496 TACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATA 6555
              |||| | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        299 TACAATCATGTCATACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATT 358

Qy       6556 GGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAG 6615
              |||||||||||||||| | ||| ||||||||||||||||||||||||||||||||||||
Db        359 GGAGCAGGAGTGCTAGGCGGAGTAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAG 418

Qy       6616 TTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTG 6675
              |||||||||||||| |||||||||||||||| |||||||| ||||||| ||  |||||
Db        419 TTCTACTACAAACTGTCTCAAGAACTAAATGGTGACATGGAATGGGTCGCTGATACCCTG 478

Qy       6676 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCT 6735
              |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        479 GTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCT 538

Qy       6736 TTAGA 6740
              ||| |
Db        539 TTAAA 543


RESULT 746
AB126146
LOCUS       AB126146                535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE5-3.
ACCESSION   AB126146
VERSION     AB126146.1  GI:47169746
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE5-3"
                     /db_xref="taxon:9542"
                     /tissue_type="thymus"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 449.7;  DB 1;  Length 535;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  499;  Conservative    4;  Mismatches   32;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGACCTCTTGTTTCCAAC 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||| ||| |||||||||||||||||||| ||||||| ||||| | |||||
Db         61 CTGGAAATAACCCGTACATCAAACCTCACCTGTGTAAATTTTAGCAGTACTATAGACACA 120
```

```
Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
             |||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
             |||||||||||||||||||||| || |||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCTCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
             | | ||||| |  |||||||| |||||||||||| |||||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
             ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
             ||||||||| |||||||||||||||||||| |||||||||||||||| ||| ||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535


RESULT 747
AB079507
LOCUS       AB079507                 540 bp    mRNA    linear   PRI 06-AUG-2002
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-7.
ACCESSION   AB079507
VERSION     AB079507.1  GI:22122081
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-7"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     repeat_region   1. .540
                     /rpt_family="HERV-W"

  Query Match             6.0%;  Score 449.4;  DB 1;  Length 540;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  480;  Conservative    4;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       6200 ACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATS 6259
             |||||||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||:
Db          2 ACAACTTCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGAACCTCTTATTTCCAATC 61

Qy       6260 TGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
```

```
             ||   ||||| ||| |||||||||||| |||||||||||||||||| |||| | ||||||
Db         62 TGAGGATAACTCATGCCTCAAACCTCAGCTGTGTAAAATTTAGCAACACTATAGACACAA 121

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
             |||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db        122 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCTCACACGAATAGTCTGCCTACCCTCAG 181

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
             |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAATAGATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTT 241

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
             ||||||  | |    | |      ||||:|||:|||||||||||||| ||||||||||||
Db        242 TGTGCTATCCCCTCATTCTTAGTGCCCCCATGACCATCTACACTGAATAAGATTTATACA 301

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
             ||||||  ||| |||||||||| ||||||||||||||||||||||||||||| ||||| ||||
Db        302 TTTATGTTATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTATGTTATTGGAG 361

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
             ||||||||||||   |||||||||||||| ||||| |||||||||||||||| |||||||
Db        362 AAGGAGTGCTAGGCAGACTAGGTACTGGCGTTGGCAGTATCACAACCTCTACCCAGTTCT 421

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
             | || |||||||||||||| |||||||| |||||||| |||| || ||||| |||||||
Db        422 AATATAAACTATCTCAAGGACTAAATGGTGACATGGAATGGGTTGCTGACTCACTGGTCA 481

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
             |||||||||| |||||||||||||||||||||||| ||||||:||||||||||||||||
Db        482 CCTTGCAAGATTAACTTAACTCCCTAGCAGCAGTATTCCTTCAAAATCGAAGAGCTTTA 540


RESULT 748
AB126143
LOCUS       AB126143                 535 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE4-1.
ACCESSION   AB126143
VERSION     AB126143.1  GI:47169743
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 535)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE4-1"
                     /db_xref="taxon:9542"
                     /tissue_type="cerebellum"
                     /note="HERV-W env"
     repeat_region   1. .535
                     /note="HERV-W env"

  Query Match             6.0%;  Score 448.7;  DB 1;  Length 535;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  498;  Conservative    4;  Mismatches   33;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||  |||||||||||||||||||| ||||||||||||
```

```
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACNTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||||||||||||||||| |||||||||||||||||||| ||||  || | |||||
Db         61 CTGGAAATAACCCATACCTCAAGCCTCACCTGTGTAAAATTTAACAATGTTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||| |||||||||||||||||||||||||||||| |||||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| ||||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTATAC 297

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| |  |||||||| |||||||||||||| |||||||||||||||||| |||
Db        298 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 356

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        357 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 416

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        417 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 476

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||| || |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        477 ACCTTGCGAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 535
```

```
RESULT 749
AB126156
LOCUS       AB126156                537 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE7-8.
ACCESSION   AB126156
VERSION     AB126156.1  GI:47169756
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 537)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE7-8"
                     /db_xref="taxon:9542"
                     /tissue_type="intestine"
                     /note="HERV-W env"
     repeat_region   1. .537
                     /note="HERV-W env"

  Query Match             6.0%;  Score 447.9;  DB 1;  Length 537;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  498;  Conservative    4;  Mismatches   35;  Indels    3;  Gaps    3;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
              |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACGGAAATAAACTCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :|||||| |||||||||||||||||| |||||||| |||||||||||||| |||||
Db         61 GTGGAAATGGCCCATACCTCAAACCTCACCTATGTAAAATGTAGCAATACTACAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              |||||||||||||||||||||||||||||||  ||||| ||||||||||||||||||||
Db        121 GCCAACTCCCAATGCATCAGGTGGGTAACTCCTTGCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||| |||||||||||||||||||| |||| |||||||||||||||||||||||||
Db        181 GGAATA-TTTTTGTCTGTGGTACCTCAGTCTATGGTTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||| ||||||:  |:|||||||||||||||| |||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGCGCCCCCTGCGACCATCTACACTGAACAAGTTTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
              | | ||||| |  |||||||| |||||||||||||| |||||||||||||||||| |||
Db        300 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 358

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
              ||||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||
Db        359 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 418

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
              |||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        419 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 478

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              ||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||||
Db        479 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 537
```

```
RESULT 750
AB120059
LOCUS       AB120059                525 bp    RNA     linear   VRL 01-FEB-2006
DEFINITION  Human endogenous retrovirus W genomic RNA, region corresponding to
            gag, isolate:HWG10-5.
ACCESSION   AB120059
VERSION     AB120059.1  GI:45330730
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
  PUBMED    15105536
REFERENCE   2  (bases 1 to 525)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30 Changjeon-dong Keum-jeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .525
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic RNA"
                     /isolate="HWG10-5"
                     /isolation_source="spleen of human"
                     /db_xref="taxon:87786"
     misc_feature    1. .525
                     /note="corresponds to gag"

  Query Match             5.9%;  Score 445.5;  DB 1;  Length 525;
```

```
  Best Local Similarity   95.8%;  Pred. No. 0;
  Matches  477;  Conservative    9;  Mismatches    9;  Indels    3;  Gaps    3;

Qy        657 RAATTGGGAMCAATTTGACCCTCAGACACTAAGAAAGAAACG-ACTTATATTCTTCTGCA 715
              :| | ||  :|||||||||||||||||||||||||||  | || ||||||||||||||||
Db         29 AATTGGGCCCCAATTTGACCCTCAGACACTAAGAAAACACCGAACTTATATTCTTCTGCA 88

Qy        716 GTGCCGCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACA-GCTAGACYT 774
              |||||||||||||||||| ||| |||||||||||||||||||||||||| |||||||:|
Db         89 GTGCCGCCTGGCACTCCTCAGGCAAGTATAAATTATAACACCATCTTACAGGCTAGACCT 148

Qy        775 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTTTTGTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAG 208

Qy        835 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGACAACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCT 268

Qy        895 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCC 954
              ||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db        269 ACCTCCCTATCCCAGCATCCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCC 328

Qy        955 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTC 388

Qy       1015 CCCAATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 1074
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        389 CCCAATTATGACCCCT-CCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCAT 447

Qy       1075 GTGCYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 1134
              ||||:||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GTGCCTTTTTCTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGAT 507

Qy       1135 AAYCCTGATGGCTATATT 1152
              ||:||||||||||||||
Db        508 AACCCTGATGGCTATATT 525


RESULT 751
CQ071133/c
LOCUS       CQ071133                559 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 6933 from Patent WO0157278.
ACCESSION   CQ071133
VERSION     CQ071133.1  GI:41041002
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 6933 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN HELA, SIGNAL = 3.9"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
              ||||| |||||||||| |||||||||:||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
              |:||||||| ||||||||||||||||||||||||||||:||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
              |||||||||||||  |||||||||||||||||||||||||:|| |||           |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
              |||||||||||| ||| ||||||||  ||||| ||||||||  || |||||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :|||||| | |||||||||| |||||||  |||||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
              |      ||    |||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
              |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 752
CQ100883/c
LOCUS       CQ100883                559 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9742 from Patent WO0157272.
ACCESSION   CQ100883
VERSION     CQ100883.1  GI:41069909
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 9742 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN PLACENTA, SIGNAL = 2.5"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
              ||||| |||||||||| |||||||||:||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500
```

```
Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:|||||||| |||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             ||||||||||||  |||||||||||||||||||||||||||||:|| |||           |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  |||  ||||||||  |||||  ||||||||  ||  |||||||||  |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||  |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  |||  ||||||  ||||||  |||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |       ||    |||||||||||||||  ||| ||||||||||||||||  |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1
```

```
RESULT 753
CQ139878/c
LOCUS       CQ139878                559 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9900 from Patent WO0157276.
ACCESSION   CQ139878
VERSION     CQ139878.1  GI:41097250
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 9900 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN BONE MARROW, SIGNAL = 3.2"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             |||||  |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:|||||||| |||||||||||||||||||||||||||||||:|||||||||||||||||||
```

```
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
             ||||||||||||  |||||||||||||||||||||||||||||:|| |||           |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||  |||  ||||||||  |||||  ||||||||  ||  |||||||||  |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||  |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  |||  ||||||  ||||||  |||||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |       ||    |||||||||||||||  ||| ||||||||||||||||  |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1
```

```
RESULT 754
CQ176105/c
LOCUS       CQ176105                559 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 7501 from Patent WO0157274.
ACCESSION   CQ176105
VERSION     CQ176105.1  GI:41170844
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 7501 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN HEART, SIGNAL = 2"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             |||||  |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:|||||||| |||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440
```

```
Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||             |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||| ||| ||||||||  ||||| ||||||||  || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | |||||||||  |||||||  |||||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |       ||    ||||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 755
CQ223232/c
LOCUS CQ223232 559 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 10071 from Patent WO0157273.
ACCESSION CQ223232
VERSION CQ223232.1 GI:41205513
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000 (03.10.00)<150> US 60/236,359<151> 27 September 2000 (27.09.00)<150> US 60/234,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170> Molecular Dynamics Sequence Listing Engine
JOURNAL Patent: WO 0157273-A 10071 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .559
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AL139245.2
EXPRESSED IN ADULT LIVER, SIGNAL = 1.9"

Query Match 5.9%; Score 445.1; DB 1; Length 559;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 492; Conservative 13; Mismatches 42; Indels 19; Gaps 2;

```
Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
             |:||||||| ||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
             |||||||||||||  |||||||||||||||||||||||:|| |||             |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
             |||||||||||| ||| ||||||||  ||||| ||||||||  || |||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | |||||||||  |||||||  |||||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
             |       ||    ||||||||||||||||  ||| ||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
             |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1
```

RESULT 756
CQ261200/c
LOCUS CQ261200 559 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 9461 from Patent WO0157277.
ACCESSION CQ261200
VERSION CQ261200.1 GI:41233680
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 9461 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .559
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AL139245.2
EXPRESSED IN FETAL LIVER, SIGNAL = 3.3"

Query Match 5.9%; Score 445.1; DB 1; Length 559;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 492; Conservative 13; Mismatches 42; Indels 19; Gaps 2;

```
Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
             ||||| |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
```

```
                |:||||||| |||||||||||||||||||||||||||||:||||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
                |||||||||||||  |||||||||||||||||||||||||||:|| |||            |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                |||||||||||  ||| ||||||||  |||||  ||||||||  ||  |||||||||  |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :|||||||  | |||||||||  |||||||  ||||||||||||||||||||||||||  |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||  ||| ||||||  ||||||  |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |        ||    ||||||||||||||  ||| ||||||||||||||||  |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                |||||||||||||:|||||||||:|||||||||:| ||||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 757
CQ298620/c
LOCUS       CQ298620                559 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9725 from Patent WO0186003.
ACCESSION   CQ298620
VERSION     CQ298620.1  GI:41259197
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 9725 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN LUNG, SIGNAL = 1.9"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
                |||||  |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
                |:||||||| |||||||||||||||||||||||||||||:|||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------C 7036
                |||||||||||||  |||||||||||||||||||||||||||:|| |||            |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                |||||||||||  ||| ||||||||  |||||  ||||||||  ||  |||||||||  |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :|||||||  | |||||||||  |||||||  ||||||||||||||||||||||||||  |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||  ||| ||||||  ||||||  |||||||||||||||||||||||: ||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |        ||    ||||||||||||||  ||| ||||||||||||||||  |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                |||||||||||||:|||||||||:|||||||||:| ||||||||||| ||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 758
CQ335348/c
LOCUS       CQ335348                559 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9442 from Patent WO0157275.
ACCESSION   CQ335348
VERSION     CQ335348.1  GI:41284419
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 9442 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN BRAIN, SIGNAL = 2.2"

  Query Match             5.9%;  Score 445.1;  DB 1;  Length 559;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  492;  Conservative   13;  Mismatches   42;  Indels   19;  Gaps    2;

Qy       6809 AATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSAGATCGAATACAACGTAKAGCAG 6868
                |||||  |||||||||| |||||||||:|||||||||:|||| | ||||||| ||:|||||
Db        559 AATCCAGAATCGTCACCGAGAAAGTTAAAGAAATTCGAGATTGCATACAACATAGAGCAG 500

Qy       6869 ARGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCC 6928
                |:||||||| |||||||||||||||||||||||||||||:|||||||||||||||||||
Db        499 AGGAGCTTCAAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCC 440

Qy       6929 CCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTR 6988
```

```
                ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||:
Db        439 CCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTATCTTTA 380

Qy       6989 ACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------C 7036
                |||||||||||||  |||||||||||||||||||||||||:|| |||           |
Db        379 ACCTCCTTGTTAAGCTTGTCTCTTCCAGAATCGAAGCTGTGAAGCTACAAATCATTCTTC 320

Qy       7037 AAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG 7096
                |||||||||||  ||| ||||||||  ||||| ||||||||  || |||||||||| |||||
Db        319 AAATGGAGCCCCAGAGGCAGTCCATAACTAAAATCTACCGTGGATCCCTGGACCAGCCTG 260

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
                :|||||||  | |||||||||  |||||||  |||||||||||||||||||||||||| |||
Db        259 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 200

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
                ||  ||| |||||| |||||| |||||||||||||||||||||||: ||||||||||||
Db        199 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 140

Qy       7217 CCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGA 7276
                |       ||    ||||||||||||||||  ||| ||||||||||||||||||| |||||
Db        139 C-------TCAACAGGGTTTTCCTGTTGAGAGAGGGTACTGAGAGACAGGACTACCTGGA 87

Qy       7277 TTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTA 7336
                |||||||||||||:|||||||||:|||||||||:| |||||||||| ||||||||||||
Db         86 TTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGGAAGGTGACCGCATCCACCTTTA 27

Qy       7337 AACACGGGGCTTGCAACTTAGYTCAC 7362
                |||||||||||||||||||:||||
Db         26 AACACGGGGCTTGCAACTTAGCTCAC 1


RESULT 759
AB097583
LOCUS       AB097583                534 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-8.
ACCESSION   AB097583
VERSION     AB097583.1  GI:31442173
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 534)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-8"
                     /db_xref="taxon:9542"
     repeat_region   1. .534
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             5.9%;  Score 444.6;  DB 1;  Length 534;
  Best Local Similarity   93.0%;  Pred. No. 0;
  Matches  502;  Conservative    4;  Mismatches   28;  Indels    6;  Gaps    4;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
                |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAATGCCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
                :|||||||||||||||||||||||||||||||||||||||||||| |||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAACAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
                |||||||||||| ||||||||||||||||||||||||||||||| |||||||||||||||||
Db        121 ACCAACTCCCAGTGCATCAGGTGGGTAACTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
                |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        181 GGAATATTTTTTGTCTGTGGTACCTCAGCCTA---TTGTTTGAATGGCTCTTCAGAATCT 237

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
                |||||||||||||||| |||||| ||||: ||:|||||||||||||||||| |||||||
Db        238 ATGTGCTTCCTCTCATTTTTAGTG-CCCCTGTGACCATCTACACTGAACAAGTTTTATAC 296

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGA 6558
                | | ||||| |  |||||||| |||||||||||||| ||||||||||||||||||| |||
Db        297 AATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCGGA 355

Qy       6559 GCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTC 6618
                ||||||||||||||  || |||||||||||||| |||||||||||| ||||||||||||
Db        356 GCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGTTC 415

Qy       6619 TACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTC 6678
                ||||||||| |||||||||||||||||| |||||||||||||| ||| ||||||||||||
Db        416 TACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGGTC 475

Qy       6679 ACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
                |||||||||| |||||||||| ||||||||||||||||||||||:||||| ||||||||||
Db        476 ACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTTTA 534


RESULT 760
AB051006
LOCUS       AB051006                541 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Homo sapiens DNA, HERV-W family, isolate:W-12-7.
ACCESSION   AB051006
VERSION     AB051006.1  GI:14020832
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S. and Lee,W.H.
  TITLE     Human endogenous retrovirus HERV-W family: chromosomal
            localization, identification, and phylogeny
  JOURNAL   AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
   PUBMED   11375061
REFERENCE   2  (bases 1 to 541)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2000) Heui-Soo Kim, Pusan National University,
            Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
            (E-mail:khs307@hyowon.cc.pusan.ac.kr, URL:http://mpl.biology.or.kr,
            Tel:82515102259, Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-12-7"
                     /db_xref="taxon:9606"
                     /chromosome="12"
     repeat_region   1. .541
                     /rpt_family="HERV-W env"

  Query Match             5.9%;  Score 444.3;  DB 1;  Length 541;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches  498;  Conservative    3;  Mismatches   35;  Indels    9;  Gaps    3;
```

```
Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||| |||||||||||||||| |||||||||||  | ||||||||||
Db          1 AACAACTTCAGCACAGTAATAAACACCACTTCCATTTTAGTAGGACTCCCTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              :||| | ||||||| || |||||| |||||||||||||||||||||||||  | | |||
Db         61 CTGGGAGTAACCCACACGTCAAACTTCACCTGTGTAAAATTTAGCAATACTTTAGAGACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              | |||||||||||| |||||||||| || |||||||||||| ||||||||||||||||||
Db        121 ATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGAATAGTCTGCCTACCCTCA 180

Qy       6379 GGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
              ||||||     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAATACTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 240

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
              |||||||||||||||||||||||||||| |||| |||:||||||||||||||||||||||
Db        241 AATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTGAACAAGATT 299

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
              |||||||| |   | || |||||||||||||||||||||||||||||||||||||||||||
Db        300 TATACAGTCA---CGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 356

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTC 6613
              | |||||||||||||||| | |||||||||||||||||| |||||||||||||| |||
Db        357 TTGGAGCAGGAGTGCTAGGCGGACTAGGTACTGGCATTGGCAGTATCACAACCTCTGCTC 416

Qy       6614 AGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCC 6673
              |||||||||||||||||||||||||||||||| |||||||| ||||  | |||||||
Db        417 AGTTCTACTACAAACTATCTCAAGAACTAAATGGTGACATGGAATGGGTTACTGACTCCC 476

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||| ||||||||||||||||||||||||||||||:||||| |||||
Db        477 TGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAG 536

Qy       6734 CTTTA 6738
              |||||
Db        537 CTTTA 541
```

```
RESULT 761
AB096035
LOCUS       AB096035                534 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-12.
ACCESSION   AB096035
VERSION     AB096035.1  GI:30698511
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 534)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-12"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .534
                     /gene="env"
     repeat_region   1. .534
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"
```

```
  Query Match             5.9%;  Score 442.4;  DB 1;  Length 534;
  Best Local Similarity   92.8%;  Pred. No. 0;
  Matches  500;  Conservative    4;  Mismatches   25;  Indels   10;  Gaps    4;

Qy       6202 AACTTCAGCACAGAAATAAACA-CCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATST 6260
              |||||||||||||||||||| | |||||| |||||||||||||||||||||||||||:|
Db          4 AACTTCAGCACAGAAATAAACATCAACTTCCATTTTAGTAGGACCTCTTGTTTCCAATCT 63

Qy       6261 GGAAATAACCCATACCTC-AAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAA 6319
              |||||||||||||||||| |||||||||||||||||||||||||      | ||||||
Db         64 GGAAATAACCCATACCTCAAAACCTCACCTGTGTAAAATTTAGCAAT-----AGACACAA 118

Qy       6320 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAG 6379
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        119 CCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTGCCCTCAG 178

Qy       6380 GAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTA 6439
              || || |||||||||||||||||||||||||||    |||||||||||||||||||||||
Db        179 GAGTACTTTTTGTCTGTGGTACCTCAGCCTATC---GTTTGAATGGCTCTTCAGAATCTA 235

Qy       6440 TGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATACA 6499
              |||||||||||||||||  |||||||:|||:|||||||||||||||||||||||||||||
Db        236 TGTGCTTCCTCTCATTCTTCATGCCCCCTATGACCATCTACACTGAACAAGATTTATACA 295

Qy       6500 GTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAG 6559
              |||||||||| |||||||| |||||||||||||||||||||||||||||||||| ||||
Db        296 ATTATGTCATACCTAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAG 355

Qy       6560 CAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCT 6619
              || |||||||||| | ||||||| |||||||||| |||||||||||||||||||||| ||
Db        356 CAAGAGTGCTAGGCGGACTAGGTTCTGGCATTGGTGGTATCACAACCTCTACTCAGTCCT 415

Qy       6620 ACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCA 6679
              |||||||||||||||||||||||||| |||||||||| ||| ||| ||||||| ||||
Db        416 ACTACAAACTATCTCAAGAACTAAATGGTGACATGGAACAGGTTGCCAACTCCCTCGTCA 475

Qy       6680 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTTTA 6738
              |||||||||||||||||||||||||||||||||||||||:||||| ||| ||||||
Db        476 CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGCGCTTTA 534
```

```
RESULT 762
AB126135
LOCUS       AB126135                539 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE1-10.
ACCESSION   AB126135
VERSION     AB126135.1  GI:47169735
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 539)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .539
```

```
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE1-10"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W env"
     repeat_region   1. .539
                     /note="HERV-W env"

  Query Match             5.9%;  Score 442.1;  DB 1;  Length 539;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  495;  Conservative    4;  Mismatches   38;  Indels    5;  Gaps    3;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||| ||| |||||||||||| |||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGTACATAAATAAACACCATTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA--TACTACATACA 6316
             :|||||||||||||||||||||||||||||||||||||||||||||  ||||||  ||||| | |||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATGTAGCAATTTACTATAGACA 120

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             |||||||||||| |||||| ||| |||||||||| |||||||||| ||||||||||||||
Db         121 CAACCAACTCCCAGTGCATCGGGTAGGTAACTCCGCCCACACAAATGGTCTGCCTACCCT 180

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             |||||  |      |||| |||||||||||||||| |||| ||||||||||||||||||||
Db         181 CAGGACAA--AAAAGTCTATGGTACCTCAGCCTATAGTTGCTTGAATGGCTCTTCAGAAT 238

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             |||||||||||||||||||||||||||| |:|||:||||||||||||||||||||||||||
Db         239 CTATGTGCTTCCTCTCATTCTTAGTGCCCTCTATGACCATCTACACTGAACAAGATTTAT 298

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| | |||| || |||||||| |||||||||||||||||||||||||||||||||||| |
Db         299 ACAATCATGTCGTACCTAAGCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTG 358

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             |||||||||||||||||  || |||||||||||||| |||||||||||| |||||||||
Db         359 GAGCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGT 418

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             |||||||||||| |||||||||||||||||||| |||||||||||||| ||| |||||||||
Db         419 TCTACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGG 478

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             |||||||||||| |||||||||| ||||||||||||||||||||:||||| ||||||||
Db         479 TCACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTT 537

Qy       6737 TA 6738
             ||
Db         538 TA 539


RESULT 763
AB097587
LOCUS       AB097587                 540 bp    DNA     linear   PRI 05-AUG-2003
DEFINITION  Macaca fuscata DNA, repeat sequence HERV-W env, isolate:JM-WE-12.
ACCESSION   AB097587
VERSION     AB097587.1  GI:31442177
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J.M., Takenaka,O. and Kim,H.S.
  TITLE     Molecular characterization and phylogenetic relationship of HERV-W
            family in the Macaca fuscata
  JOURNAL   Arch. Virol. 148 (8), 1613-1622 (2003)
   PUBMED   12898334
REFERENCE   2  (bases 1 to 540)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2002) Joo-Mi Yi, Pusan National University,
            Biology; 30, Jangjeon-dong Keumjeong-gu, Busan 609-735, Republic of
            Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JM-WE-12"
                     /db_xref="taxon:9542"
     repeat_region   1. .540
                     /note="HERV-W env"
                     /rpt_family="HERV-W"

  Query Match             5.9%;  Score 441.4;  DB 1;  Length 540;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  502;  Conservative    2;  Mismatches   34;  Indels    4;  Gaps    4;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCATTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :|||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         61 CTGGAAATAACCCATACCTCAGACCTCACCTGTGTAAAATTTAGCAATACTATAGACACA 120

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             |||||||||| |||||||||| ||| |||||||||| ||||||||| | |||||||
Db         121 ACCAACTCCCAGTGCATCAGGTGGATAAATCCTCCCACACGAATAGTCTGACCACCCTCA 180

Qy       6379 GGAATA-TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATC 6437
             | |||| |||||||||||||||| |||||| ||| ||||||||||||||||||||||||
Db         181 GAAATATTTTTTTGTCTGTGGTATCTCAGCTTATTGTTGTTTGAATGGCTCTTCAGAATC 240

Qy       6438 TATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATA 6497
             |||||||||||||||||| |||||||||| |||:|||| ||||||||||||||||||||||
Db         241 TATGTGCTTCCTCTCATTTTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATA 299

Qy       6498 CAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGG 6557
             ||  |||||||||||||| | ||||| |||||||||||||||||||||||||||||||||
Db         300 CAACTATGTCATATCTAAGGGCTGCAAC-AAAGAGTACCCATTCTTCCTTTTGTTATTGG 358

Qy       6558 AGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTT 6617
             | | |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         359 AACGGGAGTGCTAGGTGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTAGTCAGTT 418

Qy       6618 CTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGT 6677
             |||||||||||||||||||||||||||||| |||| ||||| |||||| ||||||||||
Db         419 CTACTACAAACTATCTCAAGAACTAAATGGTGACAGGGAACTGGTCGCTGACTCCCTGGT 478

Qy       6678 CA-CCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
              | ||||||||||||||||||||||||||| |||| |||||||||| ||||||||||||||
Db         479 TACCCTTGCAAGATCAACTTAACTCCCTATCAGCNGTAGTCCTTCCAAATCGAAGAGCTT 538

Qy       6737 TA 6738
             ||
Db         539 TA 540


RESULT 764
AB114965
LOCUS       AB114965                 549 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE7-7.
ACCESSION   AB114965
VERSION     AB114965.1  GI:41016473
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 549)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE7-7"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     repeat_region   1. .549
                     /note="HERV-W env"

  Query Match             5.9%;  Score 440.7;  DB 1;  Length 549;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  534;  Conservative    4;  Mismatches    2;  Indels    9;  Gaps    9;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAA-CCTCACCTGTGTAAAATTTAGCAATACTACATACAC 6317
              :|||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 CTGGAAATAACCCATACCTCAAACCCTCACCTGTGTAAAATTTAGCAATACTACATACAC 120

Qy       6318 AACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTA-CCCT 6376
             |||||| ||||||||||||||||||||||||||||||||||||||||| ||||||| ||||
Db        121 AACCAATTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGCCTGCCTACCCCT 180

Qy       6377 CAGGAATATTTTTTGTCTGTGGTA-CCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 6435
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        181 CAGGAATATTTTTTGTCTGTGGTACCCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAA 240

Qy       6436 TCTATGTGC-TTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
             ||||||||| |||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        241 TCTATGTGCTTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTT 300

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 360

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACC-TCTACTC 6613
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        361 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTTCTACTC 420

Qy       6614 AGTTCTACTACAAACTA-TCTCAAGAACTAAATGGGGACA-TGGAACGGGTCGCCGACTC 6671
             |||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db        421 AGTTCTACTACAAACTAGTCTCAAGAACTAAATGGGGACAGTGGAACGGGTCGCCGACTC 480

Qy       6672 C-CTGGTCACC-TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGA 6729
             | |||||||||| ||||||||||||||||||||||||||||||||||:|||||||
Db        481 CTCTGGTCACCTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGA 540

Qy       6730 AGAGCTTTA 6738
             |||||||||
Db        541 AGAGCTTTA 549


RESULT 765
AB096031
LOCUS       AB096031                544 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-6.
ACCESSION   AB096031
VERSION     AB096031.1  GI:30698507
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 544)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-6"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .544
                     /gene="env"
     repeat_region   1. .544
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             5.9%;  Score 438.8;  DB 1;  Length 544;
  Best Local Similarity   91.9%;  Pred. No. 0;
  Matches  501;  Conservative    3;  Mismatches   35;  Indels    6;  Gaps    4;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCC-TTTTAGTAGGACCTCTTGTTTCCAAT 59

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
              |||||| || |||||||||||| ||||| | ||||||||||||||||||   | | |||
Db         60 TTGGAAACAATCCATACCTCAAATCTCACTTATGTAAAATTTAGCAATACTGTAGATACA 119

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCT 6438
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        180 GGAATATTTTTTGTTTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCT 239

Qy       6439 ATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTATAC 6498
              |||||||||||||||||||||||||||:|||:|||||||||| ||||||||||||||||
Db        240 ATGTGCTTCCTCTCATTCTTAGTGCCCCCTATGACCATCTACACAGAACAAGATTTATAC 299

Qy       6499 AGTTATGTCATATCTAAGCCCCGCAAC--AAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
              | | ||||| || |||||||||   |  ||||||||||||||||||||||||||||||
Db        300 AATCATGTCGTACCTAAGCCCCCACAACAAAAAGAGTACCCATTCTTCCTTTTGTTATCA 359

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
              |||||||||||||||||   |||||||||||||||||| ||||||||||||||||||||
Db        360 GAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATCACAACCTCTACTCAGT 419

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGAC-TCCCTG 6675
              |||||||||||||||||||||| ||||||| |||||||||| |||| | ||| ||||||
Db        420 TCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTTCCCTG 479

Qy       6676 GTCACCTTGCAAGATCAACTT--AACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              ||||||| |||||||||||||  | ||||||||||||||||||||:||||||||||||
Db        480 GTCACCTGGCAAGATCAACTTTAACTTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAG 539

Qy       6734 CTTTA 6738
```

```
Db        540 CTTTA 544


RESULT 766
BV198554
LOCUS       BV198554                601 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm198084 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV198554
VERSION     BV198554.1  GI:48166631
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 601.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>601

  Query Match             5.8%;  Score 437.4;  DB 1;  Length 601;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches  487;  Conservative    9;  Mismatches   45;  Indels   16;  Gaps    2;

Qy       7016 GAATCGAAGCTGTRAAACTA-------CAAATGGAGCCCAAGATGCAGTCCAAGACTAAG 7068
              ||||  |||||||:||||||       |||||||||||| |||||||||||| |||||||
Db         41 GAATTAAAGCTGTAAAACTACATTCTTCAAATGGAGCCCCAGATGCAGTCCATGACTAAG 100

Qy       7069 ATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATGACATCAAA 7128
              ||||||| ||||||| |||||||||||||:||||||| | ||  |||||||||||||| ||
Db        101 ATCTACCACAGACCACTGGACCGGCCTGCTAGCCCATGCTCCAATGTTAATGACATCGAA 160

Qy       7129 GGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGA 7188
              |||||||||||| |||||||||||| ||||| |||| ||||||| ||||||||||||||||
Db        161 GGCACCCCTCCCGAGGAAATCTCAACTGCAAAACCCCTACTACACCCCAATTCAGCAGGA 220

Qy       7189 AGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAT 7248
              |||||||||||| ||:|||||||||||||||||||||||||||||| ||||||||||||||||
Db        221 AGCAGTTAGAGCAGTCGTCGGCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAT 280

Qy       7249 GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCCYTAAGCCTA 7308
              |||||||||||||||||||| ||||||||||||||||||||:|| ||||||:||||||||
Db        281 GGGGGACTGAGAGACAGGACRAGCTGGATTTCCTAGGCTGACTAGGAATCCCTAAGCCTA 340

Qy       7309 GSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGA 7368
              |:|||||||||||  | |||||||||||||| ||||||||||||||:||||| ||||
Db        341 GCTGGGAAGGTGACTGCTTCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACATCTGA 400

Qy       7369 CCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAG 7419
              ||||||         |||||||||||||||||||||||||| |||| ||||||||||||||
Db        401 CCAATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGACAAAAACAGGAGGTAAAG 460

Qy       7420 AAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAA 7479
              |||||||||||| ||:||||||:|||||||| ||  |||||||||||||| ||||||||||
Db         461 AAATAGCCAATAATCTATTGCCTGAGAGCATAGTGGGAGGGACAATGATGGGGATATAAA 520

Qy       7480 CCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGC 7539
              |||| |  ||| ||| ||||| ||| |||| ||||| |||||||||||||||||||||||||
Db         521 CCCAGGCATTCAAGCTGGCAATGGCTACCCTCTTTGAGTCCCCTCCCTTTGTATGGGAGC 580

Qy       7540 TCTGTTTTCATGCTATT 7556
              ||||| |||   |||||
Db         581 TCTGTCTTCTCCCTATT 597


RESULT 767
AX905531
LOCUS       AX905531                443 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 21394 from Patent EP1033401.
ACCESSION   AX905531
VERSION     AX905531.1  GI:40060492
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 21394 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.8%;  Score 436.8;  DB 1;  Length 443;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  435;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         12 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 62

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTKGCAACTGCACTCTTCTGGTC 122

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 182

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 242

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        243 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 302

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 362

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        363 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 422

Qy        432 TCCATTCCTTGGAATCCGTGA 452
              ||||||||||||||||||||
Db        423 TCCATTCCTTGGAATCCGTGA 443
```

```
RESULT 768
BD041064
LOCUS       BD041064                443 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD041064
VERSION     BD041064.1  GI:22582806
KEYWORDS    JP 2001269182-A/17310.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 17310 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/17310
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

  Query Match             5.8%;  Score 436.8;  DB 1;  Length 443;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  435;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         12 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 62

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTKGCAACTGCACTCTTCTGGTC 122

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 182

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 242

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||| ||||
Db        243 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 302

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        303 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 362

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        363 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 422

Qy        432 TCCATTCCTTGGAATCCGTGA 452
              |||||||||||||||||||||
Db        423 TCCATTCCTTGGAATCCGTGA 443


RESULT 769
AR741325
LOCUS       AR741325                443 bp    DNA     linear   PAT 11-OCT-2005
DEFINITION  Sequence 21394 from patent US 6783961.
ACCESSION   AR741325
VERSION     AR741325.1  GI:77435170
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Edwards,J.-B.D.M., Duclert,A. and Giordano,J.-Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: US 6783961-A 21394 31-AUG-2004;
            Genset S.A.;;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             5.8%;  Score 436.8;  DB 1;  Length 443;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches  435;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         12 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCCTTTGGGTCCCTTCCCTTTGT 62

Qy         72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTC 131
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTKGCAACTGCACTCTTCTGGTC 122

Qy        132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACC 182

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 242

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||| ||||
Db        243 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 302

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        303 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 362

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        363 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 422

Qy        432 TCCATTCCTTGGAATCCGTGA 452
              |||||||||||||||||||||
Db        423 TCCATTCCTTGGAATCCGTGA 443


RESULT 770
CQ109559
LOCUS       CQ109559                491 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 18418 from Patent WO0157272.
ACCESSION   CQ109559
VERSION     CQ109559.1  GI:41078612
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
```

```
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 18418 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL122001.5
                     EXPRESSED IN PLACENTA, SIGNAL = 1.9
                     SWISSPROT HIT: P31792, EVALUE 4.00e-26
                     EST_HUMAN HIT: AU134310.1, EVALUE 0.00e+00
                     NT HIT: AF009668.1, EVALUE 0.00e+00"

  Query Match             5.8%;  Score 433.2;  DB 1;  Length 491;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  423;  Conservative   38;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       3417 AGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCT 3476
             ||:|||||||:|||||||||||||:||:|||::|||||||||||||||:||:| ||||
Db          1 AGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCACAGAAAAAAACAGGAATAACTCT 60

Qy       3477 RGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGA 3536
             :|||||||||||||||:|||:||| |:||||||||||::||||: ||||||:||||||
Db         61 AGGAGTCCTTACACAGGTCCAAGGAACGAGCTTGCAACCCGTGGCAGACCTGAGTAAGGA 120

Qy       3537 AAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGT 3596
             ||:|||||||||||||||||||:|:||||||||||:||||| ||||||||||||| |
Db        121 AATTGATGTAGTGGCAAAGGGTTGGCCTCATTGTTTATGGGTAATGGTGGCAGTAGCAAT 180

Qy       3597 YKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCA 3656
             ::|||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||
Db        181 CTTAGTATCTGAAGCAGTTAAAATAATACAGGAAAGAGATCTTACTGTGTGGACATCTCA 240

Qy       3657 TGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACT 3716
             |||:|||||::|||||||||||||||||||||||| ||||||||||||||| |: ||||||
Db        241 TGATGTGAACAGCATACTCACTGCTAAAGGAGACCTGTGGCTGTCAGACAGCCATTTACT 300

Qy       3717 TAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCT 3776
             |||||: |||| |||||||||||||:|||||||||||:|||||||||||| ||||||||
Db        301 TAAATACCAGGGTCTATTACTTGAAGGGCCAGTGCTGCAACTGTGCACTTGGGCAACTCT 360

Qy       3777 TAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAART 3836
             ||||||||:| |||||||||||||| |||| ||||||||:||: ||||||||||| :|
Db        361 TAACCCAGCCACATTTCTTCCAGACATTGAAAAAAAGATAGAACGTAACTGTCAACAGGT 420

Qy       3837 AATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGA 3896
              ||| |||||||||||||| |||||||  |||||:||||:|||||:||||||||||:||
Db        421 GATTGCTCAAACCTATGCCGCTCGAGGAAACCTTCTAGAGGTTCCTTTGACTGATCCTGA 480

Qy       3897 CCT 3899
             |||
Db        481 CCT 483


RESULT 771
AB126132
LOCUS       AB126132                538 bp    RNA     linear   PRI 14-MAY-2004
DEFINITION  Macaca fuscata RNA transcript, HERV-W env, isolate:JM-WE1-1.
ACCESSION   AB126132
VERSION     AB126132.1  GI:47169732
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Yi,J., Takenaka,O. and Kim,H.
  TITLE     Expression and identification of HERV-W family in Japanese Monkeys
            (Macaca fuscata)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 538)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Joo-Mi Yi, Pusan National University,
            Biology;, 30, Chanjeon-dong, Keumjeong-ku, Busan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Macaca fuscata"
                     /mol_type="other RNA"
                     /isolate="JM-WE1-1"
                     /db_xref="taxon:9542"
                     /tissue_type="testis"
                     /note="HERV-W env"
     repeat_region   1. .538
                     /note="HERV-W env"

  Query Match             5.7%;  Score 425.8;  DB 1;  Length 538;
  Best Local Similarity   90.6%;  Pred. No. 0;
  Matches  491;  Conservative    4;  Mismatches   41;  Indels    6;  Gaps    4;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             |||||||||| ||| |||||||||||| ||||||||||||||||||||||||||||||||
Db          1 AACAACTTCAGTACATAAATAAACACCATTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAA--TACTACATACA 6316
             :|||||||||||||||||||||||||||||||||||||||||||||| ||||||  ||||| | |||
Db         61 CTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATGTAGCAATTTACTATAGACA 120

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCT 6376
             |||||||||||| |||||| ||| |||||||||| |||||||||| ||||||||||||
Db        121 CAACCAACTCCCAGTGCATCGGGTAGGTAACTCCGCCCACACAAATGGTCTGCCTACCCT 180

Qy       6377 CAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAAT 6436
             |||||  |     |||| |||||||||||||||| |||| ||||||||||||||||||
Db        181 CAGGACAA--AAAAGTCTATGGTACCTCAGCCTATAGTTGCTTGAATGGCTCTTCAGAAT 238

Qy       6437 CTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTTAT 6496
             |||||||||||||||||||||||| ||||||: ||:|||||||||||||||| |||||
Db        239 CTATGTGCTTCCTCTCATTCTTAGCGCCCCCTGTGACCATCTACACTGAACAAGTTTTAT 298

Qy       6497 ACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAG 6556
             ||| | ||||| |  |||||||| |||||||||||| ||||||||||||||||||| |
Db        299 ACAATCATGTCGTTCCTAAGCCCCACAACAAAAGAGTA-CCATTCTTCCTTTTGTTATCG 357

Qy       6557 GAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGT 6616
             |||||||||||||||||| || |||||||||||||| |||||||||||| |||||||||
Db        358 GAGCAGGAGTGCTAGGTGGGCTGGGTACTGGCATTGGTGGTATCACAACCCCTACTCAGT 417

Qy       6617 TCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGG 6676
             |||||||||| |||||||||||||||||| |||||||||||||| ||| |||||||||
Db        418 TCTACTACAAATTATCTCAAGAACTAAATGGTGACATGGAACGGGTTGCCAACTCCCTGG 477

Qy       6677 TCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             |||||||||||| |||||||||| |||||||||||||||||||:||||| ||||||||
Db        478 TCACCTTGCAAGGTCAACTTAAC-CCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAGCTT 536

Qy       6737 TA 6738
             ||
Db        537 TA 538


RESULT 772
AB114955
LOCUS       AB114955                549 bp    RNA     linear   PRI 01-FEB-2006
DEFINITION  Homo sapiens RNA transcript, HERV-W env, isolate:WE5-7.
ACCESSION   AB114955
VERSION     AB114955.1  GI:41016463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yi,J.M., Kim,H.M. and Kim,H.S.
  TITLE     Expression of the human endogenous retrovirus HERV-W family in
            various human tissues and cancer cells
  JOURNAL   J. Gen. Virol. 85 (PT 5), 1203-1210 (2004)
   PUBMED   15105536
REFERENCE   2  (bases 1 to 549)
  AUTHORS   Yi,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2003) Joo-Mi Yi, Pusan National University,
            Biology; 30, Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@pusan.ac.kr, Tel:82-5151-03346,
            Fax:82-5158-12962)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="other RNA"
                     /isolate="WE5-7"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     repeat_region   1. .549
                     /note="HERV-W env"

  Query Match             5.6%;  Score 423.1;  DB 1;  Length 549;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  498;  Conservative    4;  Mismatches   38;  Indels    9;  Gaps    5;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
Db          1 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACC--TCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACA 6316
Db         61 CTGGAAATAACCCATACCCTCAAACCCTCACCTGTGTAAAATTTAGCAATACGATAGACA 120

Qy       6317 CAACCAACTCCCAATGCATCAGGTGGGTAAC-TCCTCCCACACAAATAGTCTGCCTA-CC 6374
Db        121 CAACCAACTCCCAATGCATCAGGTGGGTAACTTCCCCCCCCACGAATAGTCTGCCTACCC 180

Qy       6375 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGA 6434
Db        181 CTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATTGTTGTTTGAATGGCTCTTCAGA 240

Qy       6435 ATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATTT 6494
Db        241 ATCTATGTGCTTCCTGTCATTCTTAGTGCCCCCCATGACCATCTACACTGAACAAGATTC 300

Qy       6495 ATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 6554
Db        301 ATACAATCATGTCATACCTAACCCCTGCAACAAAAGAGTACCCATTCTTCCTTTTGTTAT 360

Qy       6555 AGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCA 6614
Db        361 TGGAGCAGGAGTGCTAGGCGGAATAGGTACTGGCACTGGCGGTATCACAACCTGTACTCA 420

Qy       6615 GTTCTACTACAAACTATCTCAAGAACTAAATGGG---GACATGGAACGGGTCGCCGACT- 6670
Db        421 GTTCTACTACAAACTATTTCAAGAAATAAAGTGGTGACAAGTGGAACAGGTCGCCGAGTT 480

Qy       6671 -CCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGA 6729
Db        481 CCTCTGGTCACCGAGCAAGATCAACTTAACTTCCTAGCAGCAGTAGTCCTTCAAAATCAA 540

Qy       6730 AGAGCTTTA 6738
Db        541 AGAGTTTTA 549


RESULT 773
CQ104018
LOCUS       CQ104018                558 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 12877 from Patent WO0157272.
ACCESSION   CQ104018
VERSION     CQ104018.1  GI:41073069
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 12877 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 4.2"

  Query Match             5.6%;  Score 420.9;  DB 1;  Length 558;
  Best Local Similarity   90.3%;  Pred. No. 0;
  Matches  512;  Conservative    8;  Mismatches   36;  Indels   11;  Gaps    7;

Qy        780 GTAGAAAAGGCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCATTAAGAGACA 839
Db          1 GAAAAGAAGGCAAATGGAGTGAAGTGCCATATGTACAAACTTTCTTTTCATTAAGAGACA 60

Qy        840 ACTCACAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTC 899
Db         61 ACTCGCAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTCTCAGAGTCTACCTC 120

Qy        900 CCTATCCCAGCATCCCC--GACTCCTTCCCCAMYTAATAAGGACCCCCCTTCAACCCAAA 957
Db        121 CCTACCCCGGCATCCCCCTGACTCCTTCTCCAACTAATAAGCAACCCCCTTCAACCCAAA 180

Qy        958 TGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCC 1017
Db        181 TGGTCCAAAAGGAGATAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCC 240

Qy       1018 AATTATGACCCCTCCCAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTG 1077
Db        241 GTTTATGCCCCCT-CCAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTA 299

Qy       1078 CYTTTTYYTCTCCCAGACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAY 1137
Db        300 CCTTTTTCTCTCTCAGACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAAC 359

Qy       1138 CCTGATGGCTATATTGRTGTTTTACAAGGGTTAGGACAATTCTTTGATCTGACATGGAGA 1197
Db        360 CCTGATGGCTATATTGATG-TTTACAAGGGTTAGGACAA---TCTGATCTGACATGGAGA 415

Qy       1198 GATATATATGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 1257
Db        416 GATGTA-ATGTTACTGCTAAATCAGACACTAACCCCAAATGAGAGAAGTGCCACCATAAC 474

Qy       1258 TGCAGCCTGAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACA 1317
Db        475 TGCAGCCCGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ACAGGATGACA 532

Qy       1318 ACAGAAGGAAAGANAATGATTCCCCAC 1344
Db        533 ACAG-AGGAAAGAGAACGATTGCCCAC 558


RESULT 774
AB096037
LOCUS       AB096037                542 bp    DNA     linear   PRI 14-MAY-2003
DEFINITION  Homo sapiens DNA, repeat sequence HERV-W, isolate:SW-16.
ACCESSION   AB096037
```

```
VERSION     AB096037.1  GI:30698513
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shin,K., Yi,J. and Kim,H.
  TITLE     Isolation of Human Endogenous Retrovirus HERV-W env Fragments in
            Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 542)
  AUTHORS   Kim,H., Shin,K. and Yi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SW-16"
                     /db_xref="taxon:9606"
                     /note="Schizophrenia"
     gene            1. .542
                     /gene="env"
     repeat_region   1. .542
                     /note="human endogenous retrovirus envelope fragment"
                     /rpt_family="HERV-W"

  Query Match             5.5%;  Score 415.4;  DB 1;  Length 542;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches  493;  Conservative    3;  Mismatches   39;  Indels   12;  Gaps    5;

Qy       6199 AACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAAT 6258
             ||||||||||||||| |||||||||||||| |||||||||||  | ||||||||||
Db          1 AACAACTTCAGCACAGTAATAAACACCACTTCCATTTTAGTAGGACTCCCTGTTTCCAAT 60

Qy       6259 STGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACA 6318
             :||| | ||||||| || |||||| |||||||||||||||||||||||  | | |||
Db         61 CTGGGAGTAACCCACACATCAAACTTCACCTGTGTAAAATTTAGCAATACT-TAGAGACA 119

Qy       6319 ACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCA 6378
             | |||||||||||| |||||||||| || |||||||||| ||||||||||||||||||
Db        120 ATCAACTCCCAATGAATCAGGTGGGAAATTCCTCCCACACGAATAGTCTGCCTACCCTCA 179

Qy       6379 GGAATA-----TTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
             ||||||     |||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGAATACTTTTTTTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 239

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
             |||||||||||||||||||||||||||||| |||| |||:||||||||||||||||||||||
Db        240 AATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTGAACAAGATT 298

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
             ||||||   ||  | || ||||||||||||||||||||||||||||||||||||||||||
Db        299 TATACA---ATCACGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 355

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTC 6613
             | |||||||||||||||| | |||||||||||||||||| ||||||||||||||| |||
Db        356 TTGGAGCAGGAGTGCTAGGCGGACTAGGTACTGGCATTGGCAGTATCACAACCTCTGCTC 415

Qy       6614 AGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCC 6673
             |||||||||||||||||||||||||||||||||  |||||||| ||||  | |||||||
Db        416 AGTTCTACTACAAACTATCTCAAGAACTAAATGGTAACATGGAATGGGTTACTGACTCCC 475

Qy       6674 TGGTCACCTTGCAAGATCAAC--TTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAG 6731
             |||||||||||||| |||||  | |  |||||||||||||||||||:||||| |||
Db        476 TGGTCACCTTGCAAGTTCAACCTTAACTTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAG 535

Qy       6732 AGCTTTA 6738
             |||||||
Db        536 AGCTTTA 542


RESULT 775
CQ748604
LOCUS       CQ748604                459 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34538 from Patent WO02068579.
ACCESSION   CQ748604
VERSION     CQ748604.1  GI:42374749
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 34538 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.5%;  Score 411.9;  DB 1;  Length 459;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  414;  Conservative   29;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGR 3561
             |:|||||||||||::||| ::|||||:||||||||:|||||||||||||||||||||:
Db          1 ATGAGCTTGCAACCCGTGGTATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGG 60

Qy       3562 CYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATA 3621
             |:|||||||||:||||| |||||||||||||||::|||||||||||||||||||||||||
Db         61 CCTCATTGTTTATGGGTAATGGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATA 120

Qy       3622 ATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCT 3681
             ||||||||:|||||||||| |||||||||||||||||||:|||||::|||||||||||||||
Db        121 ATACAGGGAAGAGATCTTTCTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCT 180

Qy       3682 AAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAA 3741
             |||||||||||||||||||||||||||||:|||||||||||:|||||||||||||||||||
Db        181 AAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAA 240

Qy       3742 RGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGAC 3801
             : |||||||||||:|||| |||||||||||||||||||||:| |||||||||||||||
Db        241 GAGCCAGTGCTGCGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGAC 300

Qy       3802 AATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGA 3861
             |||||||| | || |:||:|||| |||||||||:|||| |||||||||||||||||| |
Db        301 AATGAAGAGAGGAAAGAACATAATTGTCAACAAGTAATTGCTCAAACCTATGCCACTCAA 360

Qy       3862 GGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAG 3921
             |||||||||:||||:|| ||:|||||||||| :|||| ||||||||||||||||||||||
Db        361 GGGGACCTTTTAGAGGTGCCCTTGACTGATCACGACC-TCAACTTGTATACTGATGGAAG 419

Qy       3922 TTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCA 3961
             |||||||||||||||||||||| |||||: |||||||||
Db        420 TTCCTTTGTAGAAAAAGGACTTCAAAAAGCAGGGTATGCA 459


RESULT 776
CQ750414
LOCUS       CQ750414                492 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 36348 from Patent WO02068579.
ACCESSION   CQ750414
VERSION     CQ750414.1  GI:42381566
KEYWORDS    .
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 36348 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.5%;  Score 411.1;  DB 1;  Length 492;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  420;  Conservative   43;  Mismatches   22;  Indels    3;  Gaps    2;

Qy       3269 TCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAG 3328
              | |||||||:  ||:|||||||||| :|||||:|||||:|||||||||||||||||||||
Db             8 TTCCCAGGTACAGCAAAATAGCCAGACCATTATATACACTAATTAAGGAAACTCAGAAAG 67

Qy       3329 CCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAG 3388
              ||||||||||||||:|||||||||:|:||||||:::|||  ||||||||||||  ||
Db            68 CCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAAGCAGCTTTCCAGGCCCTAAAGA 127

Qy       3389 AAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATA 3448
              | | |||  ||||||||:||||||||||||:||||||||:| ||||||||||||:|||:|||
Db           128 AGGCCCT--AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGAGCAAGACTTTTCTTTATA 185

Qy       3449 YRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCT 3508
              ::||| || |||||||||:||:|||||||:|||||||||||||||:|||:|||||: ||||
Db           186 TGTCATAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTACACAGATCCAAGGGACAAGCT 244

Qy       3509 TGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATT 3568
              |||||||::||||::||||||:||||||||:|||||||||||||||||||:|:|||||
Db           245 TGCAACCCATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATT 304

Qy       3569 GTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGG 3628
              |||||:|||||| | ||||||||||||::|||||||||||||||||||||||||||||||
Db           305 GTTTACGGGTAGTAGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAGG 364

Qy       3629 GRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAG 3688
              |:|||||| ||| |||||||||||| ||||:|||||::|||||| ||||||||||||
Db           365 GAAGAGATTTTATTGTGTGGACATCTTATGATGTGAATGGCATACTGACTGCTAAAGGAG 424

Qy       3689 ACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAG 3748
              ||||||||||||||||||:||||||||||:||||||||||||||||||:||||||
Db           425 ACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAG 484

Qy       3749 TGCTGCRA 3756
              ||||| :|
Db           485 TGCTGTGA 492


RESULT 777
CQ050965
LOCUS       CQ050965                488 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 1785 from Patent WO0157270.
ACCESSION   CQ050965
VERSION     CQ050965.1  GI:41025471
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 1785 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HBL100, SIGNAL = 1.8"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db             1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
              ||||||||||||  |||||||||||||||||||||||  ||| ||||||| |||||||| |
Db            61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
              | |||||||||  |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db           121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
              |||||||||||||||:||||||||||||||||||||||||||||||||||| |||||||
Db           181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
              |||||||||| |||||| |||||||||||||||||||| || |||||||| ||||||
Db           241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
              ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||||
Db           300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | |||||||||||||||||||||||  | ||||| |||||||||| |||
Db           360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db           420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | |||||||
Db           480 CTCTGAGTC 488


RESULT 778
CQ065989
LOCUS       CQ065989                488 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 1789 from Patent WO0157278.
ACCESSION   CQ065989
VERSION     CQ065989.1  GI:41035858
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 1789 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HELA, SIGNAL = 1.7"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||||  |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||  |||||| ||||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             |||||  |||||||| |||||||  |||||||||||  |||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             ||||  ||  |  |||||||||||||||||||||||||  |  |||||  |||||||||||  |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||  |||||||||||||||||||||||  |||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 779
CQ093006
LOCUS       CQ093006                488 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1865 from Patent WO0157272.
ACCESSION   CQ093006
VERSION     CQ093006.1  GI:41062032
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 1865 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .488
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC004897.1
                      EXPRESSED IN PLACENTA, SIGNAL = 10"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 1576
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
             ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||||  |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
             | ||||||||| |||||||||||||||||||| ||||||| ||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
             |||||||||||||||:|||||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
             |||||||||||  |||||| ||||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
             |||||  |||||||| |||||||  |||||||||||  |||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
             ||||  ||  |  |||||||||||||||||||||||||  |  |||||  |||||||||||  |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
             |||||||||||||  |||||||||||||||||||||||  |||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 780
CQ131815
LOCUS       CQ131815                488 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1837 from Patent WO0157276.
ACCESSION   CQ131815
VERSION     CQ131815.1  GI:41089171
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 1837 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .488
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC004897.1
                      EXPRESSED IN BONE MARROW, SIGNAL = 1.5"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
```

```
                 ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
                 ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
                 | |||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db         121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
                 ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db         181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
                 |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db         241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
                 |||||  |||||||| |||||| ||||||||||  |||||||||||||||||||||||
Db         300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
                 ||||  ||  | |||||||||||||||||||||||  | ||||| |||||||||| |||
Db         360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
                 |||||||||||||  |||||||||||||||||||||  |||||||||||||||||||||||
Db         420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                 | ||||||
Db         480 CTCTGAGTC 488


RESULT 781
CQ170381
LOCUS       CQ170381                488 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1777 from Patent WO0157274.
ACCESSION   CQ170381
VERSION     CQ170381.1  GI:41165117
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 1777 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HEART, SIGNAL = 1.2"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
                 ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
                 ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
                 | |||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db         121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
                 ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db         181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
                 |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db         241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
                 |||||  |||||||| |||||| ||||||||||  |||||||||||||||||||||||
Db         300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
                 ||||  ||  | |||||||||||||||||||||||  | ||||| |||||||||| |||
Db         360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
                 |||||||||||||  |||||||||||||||||||||  |||||||||||||||||||||||
Db         420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                 | ||||||
Db         480 CTCTGAGTC 488


RESULT 782
CQ199555
LOCUS       CQ199555                488 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1832 from Patent WO0157271.
ACCESSION   CQ199555
VERSION     CQ199555.1  GI:41185219
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 1832 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN BT474, SIGNAL = 0.99"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
                 ||||||||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
                 ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
                 | |||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||||
Db         121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
```

```
                 ||||||||||||||||:||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
                ||||||||||| |||||| |||||||||||||||||||| || |||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
                ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
                |||| || | |||||||||||||||||||||||| | ||||| |||||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
                |||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 783
CQ215005
LOCUS       CQ215005                488 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1844 from Patent WO0157273.
ACCESSION   CQ215005
VERSION     CQ215005.1  GI:41196150
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 1844 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.8"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
                ||||||||||||||||||||||||||||||||||| |||||||||||  | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
                ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
                | |||||||| |||||||||||||||||| ||||||| ||||||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
                ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
                |||||||||||  |||||| |||||||||||||||||||| || |||||||||| ||||||
Db         241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
                ||||| |||||||| ||||||| |||||||||||  |||||||||||||||||||||||
Db         300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
                |||| || | |||||||||||||||||||||||| | ||||| |||||||||||| |||
Db         360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
                |||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||
Db         420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db         480 CTCTGAGTC 488


RESULT 784
CQ253598
LOCUS       CQ253598                488 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 1859 from Patent WO0157277.
ACCESSION   CQ253598
VERSION     CQ253598.1  GI:41226076
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 1859 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.4"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
                ||||||||||||||||||||||||||||||||||| |||||||||||  | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
                ||||||||||||  |||||||||||||||||||||||  ||| ||||||| ||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
                | |||||||| |||||||||||||||||| ||||||| ||||||||||||||||||||
Db         121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
                ||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db         181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
                |||||||||||  |||||| |||||||||||||||||||| || |||||||||| ||||||
```

```
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
              ||||| |||||||| ||||||| |||||||||| ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | |||||||||||||||||||||||||| | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 785
CQ290692
LOCUS       CQ290692                488 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 1797 from Patent WO0186003.
ACCESSION   CQ290692
VERSION     CQ290692.1  GI:41251269
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 1797 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN LUNG, SIGNAL = 1.6"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
              |||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
              |||||||||||| |||||||||||||||||||||| ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
              | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
              |||||||||||||||:||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
              |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
              ||||| |||||||| ||||||| |||||||||| ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | |||||||||||||||||||||||||| | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
                | ||||||
Db        480 CTCTGAGTC 488


RESULT 786
CQ327720
LOCUS       CQ327720                488 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 1814 from Patent WO0157275.
ACCESSION   CQ327720
VERSION     CQ327720.1  GI:41276287
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 1814 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN BRAIN, SIGNAL = 0.91"

  Query Match             5.5%;  Score 409.6;  DB 1;  Length 488;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  451;  Conservative    1;  Mismatches   35;  Indels    2;  Gaps    2;

Qy       1517 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCGTGCCTCTCTG 1576
              |||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db          1 GAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGTCATTGAGGAAGCATACCTCTCTG 60

Qy       1577 TCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCT 1636
              |||||||||||| |||||||||||||||||||||| ||| ||||||| |||||||| |
Db         61 TCACCTGACTCTATTGAAGGCCAACTAATCTTAAAGGATAAATTTATCAGTCAGTCAGTT 120

Qy       1637 GCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAACC 1695
              | ||||||||| |||||||||||||||||||||| ||||||| ||||||||||||||||
Db        121 GTAGACATTAGAAAAAAACTTCAAAAGTCTGCCTTAGGCCCAGAGCAAAACTTAGAAACC 180

Qy       1696 CTATTGAACTTGGCAACYTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAG 1755
              |||||||||||||||:||||||||||||||||||||||||||||||||| |||||||
Db        181 CTATTGAACTTGGCAACCTCGGTTTTTTATAATAGAGATCAGGAGGAGCAGGTGGAACAG 240

Qy       1756 GACAAACGGGATTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTT 1815
              |||||||||| |||||| |||||||||||||||||||| || ||||||||| ||||||
Db        241 GACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAGTCATGGCCTTCAGGCAAGCGGACTT 299

Qy       1816 TGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGAATGCCTAATAGGGCTTGCTTCCA 1875
              ||||| |||||||| ||||||| |||||||||| ||||||||||||||||||||||||
Db        300 TGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATCAAATGCCTAATAGGGCTTGCTTCCA 359

Qy       1876 GTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCCTTCG 1935
              |||| || | |||||||||||||||||||||||||| | ||||| |||||||||| |||
Db        360 GTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTCAGAATAGAAATAAGCCGCCCCCTCG 419

Qy       1936 TCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGTC 1995
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
```

```
Db        420 TCCATGCCCCTTATGTCAAGGGAATCACTGGAAGGTCCACTGCCCCAGGGGACAAAGGTC 479

Qy       1996 TTTTGAGTC 2004
              | ||||||
Db        480 CTCTGAGTC 488


RESULT 787
CQ099994/c
LOCUS       CQ099994                504 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 8853 from Patent WO0157272.
ACCESSION   CQ099994
VERSION     CQ099994.1  GI:41069020
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 8853 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN PLACENTA, SIGNAL = 4.6"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
              ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
              |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAA 7038
              |||||| |||| || |||||||||||||||||||||||||||:|| |||            |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
              ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
              |||||| | ||||||||| |||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
              ||| |||||| |||||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              ||||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              ||||||||  ||:   || |||:|||||||||:|||||||||||||| |||||||||||||||
Db         84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||:||||
Db         24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 788
CQ138977/c
LOCUS       CQ138977                504 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 8999 from Patent WO0157276.
ACCESSION   CQ138977
VERSION     CQ138977.1  GI:41096349
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 8999 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN BONE MARROW, SIGNAL = 2.5"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
              ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
              |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA------------CAA 7038
              |||||| |||| || |||||||||||||||||||||||||||:|| |||            |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
              ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
              |||||| | ||||||||| |||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
              ||| |||||| |||||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
              ||||||||||||  |||||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
              ||||||||  ||:   || |||:|||||||||:|||||||||||||| |||||||||||||||
Db         84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
              |||||||||||||||||:||||
Db         24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 789
CQ175631/c
LOCUS       CQ175631                504 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 7027 from Patent WO0157274.
ACCESSION   CQ175631
```

```
VERSION     CQ175631.1  GI:41170370
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 7027 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN HEART, SIGNAL = 2.2"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
             |||||| |||| || ||||||||||||||||||||| ||:|| |||          |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| ||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 790
CQ222372/c
LOCUS       CQ222372                 504 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9211 from Patent WO0157273.
ACCESSION   CQ222372
VERSION     CQ222372.1  GI:41204496
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 9211 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.8"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
             |||||| |||| || ||||||||||||||||||||| ||:|| |||          |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | ||||||||| ||||||   ||| ||||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| ||||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 791
CQ260325/c
LOCUS       CQ260325                 504 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 8586 from Patent WO0157277.
ACCESSION   CQ260325
VERSION     CQ260325.1  GI:41232805
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 8586 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN FETAL LIVER, SIGNAL = 2.5"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||           |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             |||||||| |||||||||||| |||||| ||||||| ||||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | |||||||||| ||||||   ||| |||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 792
CQ298000/c
LOCUS       CQ298000                 504 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 9105 from Patent WO0186003.
ACCESSION   CQ298000
VERSION     CQ298000.1  GI:41258577
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 9105 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN LUNG, SIGNAL = 2.4"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  ||||||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA-----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||           |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             |||||||| |||||||||||| |||||| ||||||| ||||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | |||||||||| ||||||   ||| |||||||| |||||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
                ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||| |||||||||||||| || |||||||||| ||||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||||| ||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 793
CQ334456/c
LOCUS       CQ334456                 504 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 8550 from Patent WO0157275.
ACCESSION   CQ334456
VERSION     CQ334456.1  GI:41283488
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 8550 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
```

```
                /db_xref="taxon:9606"
                /note="MAP TO AL049870.3
                EXPRESSED IN BRAIN, SIGNAL = 2.3"

  Query Match             5.3%;  Score 400.4;  DB 1;  Length 504;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  438;  Conservative    9;  Mismatches   45;  Indels   12;  Gaps    1;

Qy       6871 GAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCRATGGATGCCCTGGATTCTCCCC 6930
             ||||||| || |||  |||||||||||||||||||||: ||||||||||||||||||||
Db        504 GAGCTTCAAAGCACCAGACCCTGGGGCCTCCTCAGCCAGTGGATGCCCTGGATTCTCCCC 445

Qy       6931 TTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRAC 6990
             ||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||:||
Db        444 TTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAAC 385

Qy       6991 CTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTA----------CAA 7038
             |||||| |||| || |||||||||||||||||||| ||:|| |||           |||
Db        384 CTCCTTATTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGGCTCTTCAA 325

Qy       7039 ATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYT 7098
             ||||||||| |||||||||||| |||||| ||||||| |||||||||||||| |||||:|
Db        324 ATGGAGCCCCAGATGCAGTCCATGACTAAAATCTACCACAGACCCCTGGACCAGCCTGCT 265

Qy       7099 AGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCA 7158
             |||||| | |||||||| |||||   ||| ||||||||| |||||||||| |||||
Db        264 AGCCCATGCTCTGATGTTGATGACACTGAAGTCACCCCTCCCGAGGAAATCTCAACTGCA 205

Qy       7159 CAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCC 7218
               ||| |||||| |||||||||||||||||||||||||||||| ||: || ||||||| ||
Db        204 TGACCCCTACTATGCCCCAATTCAGCAGGAAGCAGTTAGAGCAGTCATCAGCCAACCACC 145

Qy       7219 CCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATT 7278
             ||||||||||||  |||||||||||||| || |||||||||| ||||||||||||||||
Db        144 CCAACAGCACTTGAGTTTTCCTGTTGAGAGGGCGGACTGAGAGAAAGGACTAGCTGGATT 85

Qy       7279 TCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAA 7338
             |||||||| ||:   || |||:|||||||||:|||||||||||| |||||||||||||||
Db          84 TCCTAGGCCGACAGCGAGTCCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAA 25

Qy       7339 CACGGGGCTTGCAACTTAGYTCAC 7362
             |||||||||||||||||||:||||
Db          24 CACGGGGCTTGCAACTTAGCTCAC 1


RESULT 794
CQ744629
LOCUS       CQ744629                480 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30563 from Patent WO02068579.
ACCESSION   CQ744629
VERSION     CQ744629.1  GI:42360762
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 30563 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.3%;  Score 396.9;  DB 1;  Length 480;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  410;  Conservative   42;  Mismatches   28;  Indels    3;  Gaps    2;

Qy       3274 AGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAAT 3333
             | | :  ||:||||||||||||: ||||:| |||:|||||||||||||||||||||||| ||
Db          1 ATGAAAAGCGAAATAGCCAGGCTATTATACACACTAATTAAGGAAACTCAGAAAGCCGAT 60

Qy       3334 ACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAAGGC 3393
             |||||||||:|||||||||:|: ||||||:::||| |||||||||||||||  ||  | | |
Db         61 ACCCATTTAGTAAGATGGACACATGAAGCAGAAGCGGCTTTCCAGGCCCTAAAGAAGGCC 120

Qy       3394 CTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYRTCA 3453
             ||  ||||||||:|||||||| ||||:||||||||:|||||||||||||||:||:|||::|||
Db         121 CT--AACCCAAGCCCCAGTGGTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATGTCA 178

Qy       3454 CAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTGCAA 3513
             ||| || ||||||:||:|||||||:||||||||| |||||:|||:|||||: |||||||||
Db         179 CAG-AAGAAACAGGAATAGCTCTAGGAGTCCTTTCACAGGTCCGAGGGACCAGCTTGCAA 237

Qy       3514 CCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTA 3573
             ||::||||::||||||:||||||||:|||||||||||||||||||  :||||||||||
Db         238 CCCGTGGCATACCTGAGTAAGGAAACTGATGTAGTGGCAAAGGGTTGTTCTCATTGTTTA 297

Qy       3574 YGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGA 3633
             :|||||||||  ||||||||||::|||||| |||||||||||||||||||||||:|||
Db         298 TGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCAGAAGCAGTTAAAATAATACAGGGAAGA 357

Qy       3634 GATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTG 3693
             |||||||||||||||||||||||||:|||||::|||||||||||||||||| |||||||
Db         358 GATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGAAGACTTG 417

Qy       3694 TGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTG 3753
             |||||||||||||||: ||||||||||:|||||||| ||||||||||:||||||||||
Db         418 TGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCCATTACTTGAAGGGCCAGTGCTG 477

Qy       3754 CRA 3756
                :|
Db         478 TAA 480


RESULT 795
AF328493/c
LOCUS       AF328493                534 bp    DNA     linear   PRI 11-JAN-2001
DEFINITION  Homo sapiens clone 4ptel_2038Br sequence.
ACCESSION   AF328493
VERSION     AF328493.1  GI:12083133
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Riethman,H.C. and Moyzis,R.K.
  TITLE     Integration of telomeric DNA sequences with the human reference
            sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 534)
  AUTHORS   Riethman,H.C. and Moyzis,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2000) Molecular Genetics, The Wistar Institute,
            3601 Spruce St., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4ptel"
                     /clone="4ptel_2038Br"
                     /note="half-YAC yRM2038"

  Query Match             5.1%;  Score 385.7;  DB 1;  Length 534;
  Best Local Similarity   85.4%;  Pred. No. 0;
  Matches  450;  Conservative    1;  Mismatches   69;  Indels    7;  Gaps    2;
```

```
Qy       1668 CGTAGGCCCGGAGCAAAACTTAGAAACCCTATTGAACTTGGCAACYTCGGTTTTTTATAA 1727
Db        523 CCTTAGGCCGGAGCAAAAACTTAGAAACCTATTGAACTTGGCAGCCTTGGTTTTTTATAA 464

Qy       1728 TAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTT 1787
Db        463 TAGAGATCAAGAGGAGCAGGCAGAATGGGATAAATGGGATAAGAAAAAGGCCCAC----T 408

Qy       1788 TAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAA 1847
Db        407 TAGTCATGGCCTTCAGGCAAGTGGACTTTGGAGGCTCTGGAACAGGGAAAGGCTGGGCAA 348

Qy       1848 ATTGAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATT 1907
Db        347 ATCAAATGCCTAATAGGGGCTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATT 288

Qy       1908 GTCCAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGA 1967
Db        287 GTCTGAATAGAAATAATCCCCACCCCCTTGTCCATGCCCTTATGTCAAGGGAATCACTGTA 228

Qy       1968 AGGCCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCC 2027
Db        227 AGGCCCACTGCCCCAGGGGATGAAGGTTCTCTTAGTCAGAAGCCACTAACCAGATGATCC 168

Qy       2028 AGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCACCCTCACAGAGC 2087
Db        167 AGCAGCAGGACTGAGGGTGCCTGGGGCAAGTGCCAGCCCATGCCATCACCCTCACTGGGC 108

Qy       2088 CCTGGGTATGCTTGACCATTGAGGGCCAGGAAGGT---TGTCTCCTGGACACTGGTGCGG 2144
Db        107 CCCGGGTATGCTTGACCATTGAGGACCAGGAGGTTAACTGCCTCCTGGACACTGGTGCAG 48

Qy       2145 TCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCCTCCAGATCT 2191
Db         47 CCTTCTCAGTCTTACTCTCCTCTCCCGGACAACTGTCCTCCAGATCT 1
```

RESULT 796
CQ752562

```
LOCUS       CQ752562                444 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 38496 from Patent WO02068579.
ACCESSION   CQ752562
VERSION     CQ752562.1  GI:42389067
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo..
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 38496 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.1%;  Score 384.5;  DB 1;  Length 444;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  392;  Conservative   28;  Mismatches   24;  Indels    1;  Gaps    1;
```

```
Qy       3502 AYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGR 3561
Db          1 ATGAACTTGCCAACTGTGGCATACCTGAGTAAGGAAATTGATGCAGTGGCAAAGGGTTGG 60

Qy       3562 CYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATA 3621
Db         61 CCTCATTGTTTATGGGTAGTGGCAGCGGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATA 120

Qy       3622 ATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCT 3681
Db        121 ATACATGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCT 180

Qy       3682 AAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAA 3741
Db        181 AAAGGAGACTTGTGGCTGTCAGACGATCGTTTACTTAAACATCAGGCTCTATTACTTGAA 240

Qy       3742 RGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGAC 3801
Db        241 GGGCCTGTGCTGCAACTGCGCGCCTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGAC 300

Qy       3802 AATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGA 3861
Db        301 AATGAAGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACCCCGCTCGA 360

Qy       3862 GGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAG 3921
Db        361 GGGGACCTTTTAGAGGTTCCCTTGACTGATTCCGACCTT-AACTTGTATACTGATGGAAG 419

Qy       3922 TTCCTTTGTAGAAAAAGGACTTCGA 3946
Db        420 TTCCTTTGTAGAAAAAGGACCTTGA 444
```

RESULT 797
AR344428

```
LOCUS       AR344428                438 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 201 from patent US 6582703.
ACCESSION   AR344428
VERSION     AR344428.1  GI:33740369
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 201 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="unknown"
                     /mol_type="genomic DNA"

  Query Match             5.1%;  Score 382.5;  DB 1;  Length 438;
  Best Local Similarity   79.9%;  Pred. No. 0;
  Matches  350;  Conservative   82;  Mismatches    5;  Indels    1;  Gaps    1;
```

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTYCYTCRGTAYRKGGATGAYTTAMTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
Db         61 WTAGCCACCCATTCAGAAACCTTGTGSCAYCAAGCCACCCAAGYRCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACMAARGGCTCASCTCTGCTCACASCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
Db        181 TACTTAGGGCTAAAATTATCCAAAGKCRCCAGRRCCCTCAGWGAGGAACGTATCCAGCST 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
Db        241 ATACTGGVTTATCCYCATCCCAWAACCMTAAAGCAACTAASARSGTTCCTTGGCATAWCA 300
```

```
Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              |:||||||||:||:|||||||  |||:|:|:  |:||::|||||||::||||||:||||::
Db        301 GSYTTCTGCCRAATATGGATT-CCCVGRTACAGYRARRTAGCCAGRCCATTAWRTACAYK 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||:|||||||:|||||||:||:|||||||||:|:||||:::::|||| 
Db        360 ADYTAAGGAAACTCARAAAGCCARTACCCATWTAGTAAGATGGACAYCTGARRCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 798
AR344405
LOCUS       AR344405                438 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 128 from patent US 6582703.
ACCESSION   AR344405
VERSION     AR344405.1  GI:33740346
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 128 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  375;  Conservative   45;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||:||||||||||:|||  |||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTAAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||||||:||||||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:||||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||  ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:| 
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| |||||||:  ||:|::||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAARRTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| |||||||:||||||||||:|||||||||:|:||||||:::|||| 
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACAYCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 799
AR344406
LOCUS       AR344406                438 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 129 from patent US 6582703.
ACCESSION   AR344406
VERSION     AR344406.1  GI:33740347
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 129 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             5.1%;  Score 381.3;  DB 1;  Length 438;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  378;  Conservative   41;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||||||:|||  |||||||||||||||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||||||:||||:|||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||||:|||||||||||:|:|:||||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||  ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:| 
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| |||||||:  ||:| |||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAGATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| |||||||:||||||||||:|||||||||:|:||||||:::|||| 
Db        360 AATTAAGGAAACTCAAAAAGCCARTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437


RESULT 800
AR344408
LOCUS       AR344408                438 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 131 from patent US 6582703.
ACCESSION   AR344408
VERSION     AR344408.1  GI:33740349
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
```

Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
TITLE Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting
JOURNAL Patent: US 6582703-A 131 24-JUN-2003;
Bio Merieux; Marcy l'Etoile;
FRX;

FEATURES Location/Qualifiers
source 1. .438
/organism="unknown"
/mol_type="mRNA"

Query Match 5.1%; Score 381.3; DB 1; Length 438;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 378; Conservative 41; Mismatches 18; Indels 1; Gaps 1;

```
Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||:||||||||||:|||  ||||||||||||||
Db          1 GACTTGAGCCAGTCYTCATACCTGGACAYTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
             ||:||: ||::|||||||||||||||||||||||||||||||:|||||:||||||||
Db         61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
             ||:|||||||||||||:||||||||||:|:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
             ||||||||:||||:|||||||||||||||||:  |||||||||||||:::|||||||||
Db        181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
             ||||||| ||||||||||:|||||||||||||||||||| :| :|||||||||:|||:|
Db        241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
             ||:||||||| ||:||||||| |||||||:  ||:|||||||||| :|||||:|||||:
Db        301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAATAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
             ||||||||||||||| ||||||||||||||||||:|||||||||:| ||||||:::||||
Db        360 AATTAAGGAAACTCAAAAAGCCAATACCCATTTAGTAAGATGGACATCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
             ||||||||||||||  ||
Db        420 GGCTTTCCAGGCCCTAAA 437
```

RESULT 801
AB024472
LOCUS AB024472 429 bp DNA linear PRI 21-OCT-1999
DEFINITION Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV, isolate:ORW3-6.
ACCESSION AB024472
VERSION AB024472.1 GI:6063503
KEYWORDS .
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (sites)
AUTHORS Kim,H.S., Takenaka,O. and Crow,T.J.
TITLE Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates
JOURNAL J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
PUBMED 10573154
REFERENCE 2 (bases 1 to 429)
AUTHORS Kim,H.-S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital, Psychiatry; Headington, Oxford OX3 7JX, U.K.
(E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397, Fax:441865244990)
FEATURES Location/Qualifiers
source 1. .429
/organism="Pongo pygmaeus"
/mol_type="genomic DNA"
/isolate="ORW3-6"
/db_xref="taxon:9600"
/sex="Male"
/tissue_type="Blood"
misc_feature <1. .>429
/note="similar to pol gene of human endogenous retrovirus family (HERV-W) and multiple sclerosis-associated retrovirus (MSRV)"

Query Match 5.1%; Score 380.9; DB 1; Length 429;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 386; Conservative 25; Mismatches 18; Indels 1; Gaps 1;

```
Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||| :||| |||||||||||| |||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGTGACTATGCACTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||  |||| ||||||||:||||| |||||||||||||||| |||||||
Db         61 AGAAAAGATAGAAAGTAACAGTCAACAAGTAATTGCTCAAACCTATGCCACTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:|||||||||||: ||| |||||||||||||||||||||||||||
Db        121 CCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             ||||||| |||||||| ||||||:|||||||||||||| | ||||||||||||||:|||
Db        180 TTGTAGAACAAGGACTTTGAAAAGTGGGGTATGCAGTGATTAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             ||||||||||||||||||||||||||||||||:||||:||||||||||||||:|||:|
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             |||||||||||||||||||:||||||||:|||||||||||| |||||||:|:||||||:
Db        300 AGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATATAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| ||||||||||::|||||||:|:|||||||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGG 419

Qy       4167 GAACACCTAT 4176
             ||||||||||
Db        420 GAACACCTAT 429
```

RESULT 802
AB024473
LOCUS AB024473 429 bp DNA linear PRI 21-OCT-1999
DEFINITION Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV, isolate:GIW4-1.
ACCESSION AB024473
VERSION AB024473.1 GI:6063504
KEYWORDS .
SOURCE Hylobates agilis (agile gibbon)
ORGANISM Hylobates agilis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (sites)
AUTHORS Kim,H.S., Takenaka,O. and Crow,T.J.
TITLE Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates
JOURNAL J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
PUBMED 10573154
REFERENCE 2 (bases 1 to 429)
AUTHORS Kim,H.-S.
TITLE Direct Submission

```
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-1"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.1%;  Score 380.5;  DB 1;  Length 429;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  385;  Conservative   26;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:||||  ||||||||||||||||||||||:| ||||||||||||||||||
Db             1 AGTGCTGCGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||||: |||| ||||||||||||  ||| ||||||
Db            61 AGAAAAGATAGAACATAACTGTCAACAAGCAATTGCTCAAACCTATGCTGCTCTAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||:|||| ||||||||||||||||||||||||||||
Db             121 CCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||||| |||||: |||||||||||||||||||| ||||||||:|||
Db             180 TTGTAGAAAAAGGACTTCAAAAAGCAGGGTATGCAGTGGTCAGTGATTATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||| ||||||||||||:||||:||||||||||||||:|||:|
Db             240 AAAGTAATCCCCTCACTCTAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTA 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              ||||| ||||||||||||||:||||||||:||||||| ||||||||||:|:||||||:
Db             300 GGGCAATAGAATTAGGAGAAGGAAAAAGGGTAAATATAAATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| ||||||||||::|||||||:|:|||||||||||||||||||: ||:
Db             360 ACCTAGTCCTCCATGCCCATGCGGCAATATGGAGAGAAAGGGAATTCCTAACTTCCAAGG 419

Qy       4167 GAACACCTAT 4176
              |||||||||
Db             420 AAACACCTAT 429


RESULT 803
AR344407
LOCUS       AR344407                438 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 130 from patent US 6582703.
ACCESSION   AR344407
VERSION     AR344407.1  GI:33740348
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 130 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             5.1%;  Score 379.7;  DB 1;  Length 438;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  377;  Conservative   41;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||||||||||||:|||  |||||||||||||
Db             1 GACTTGAGCCAGTCCTCATACCTGGACACTCTTGTCCTTCGGTACATGGATGATTTACTT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              ||:||: ||::|||||||||||||||||||||||||||||||:|||||:||||||||
Db            61 TTAGCCACCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAGCACTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||||:||||||||||:|:|:|||||:|||||||||||||||||||||
Db             121 GCTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 180

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:||||||||||||||||:  |||||||||||||:::|||||||||
Db             181 TACTTAGGGCTAAAATTATCCAAAGGCACCAGAACCCTCAGTGAGGAACGTATCCAGCCT 240

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| |||||||||||:|||||||||||||||||||| :| :||||||||:|||:|
Db             241 ATACTGGGTTATCCTCATCCCAAAACCCTAAAGCAACTAACAGCGTTCCTTGGCATAACA 300

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              ||:|||||| ||:|||||| |||||||:  ||:|| ||||||| :|||||:|||||:
Db             301 GGTTTCTGCCAAATATGGATT-CCCAGGTACAGCAAAGTAGCCAGACCATTAAATACACG 359

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||| ||||||| |||||||||:|||||||:|:||||||:::||||
Db             360 AATTAAGGAAACTCAAAAAGCCAGTACCCATTTAGTAAGATGGACACCTGAAGCAGAAGT 419

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||| ||
Db             420 GGCTTTCCAGGCCCTAAA 437


RESULT 804
AB024469
LOCUS       AB024469                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV,
            isolate:ORW3-3.
ACCESSION   AB024469
VERSION     AB024469.1  GI:6063500
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="ORW3-3"
                     /db_xref="taxon:9600"
```

```
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.1%;  Score 378.9;  DB 1;  Length 429;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  384;  Conservative   26;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||| :||||  |||||||||||||||||||||||:| ||||||||||||||||||||
Db             1 AGTGCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||| ||||||:||:|||||||||||||:||||| |||||||||| | |||| |||||||
Db            61 AGATAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACGTCACTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||:|||| |||||||||||| ||||||||||||||
Db           121 CCTTTTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATATTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| |||||:||||||| ||||||||||||||||||||||:|||
Db           180 TTGTAGAAAAAGGACTTCAAAAAGTGGGGTATCCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||||||||||||||||||||||||||||:||||:|||| ||| |||||||:|||:|
Db           240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGGACTCATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||||||||||||||:||||||||:|||||||||||||||||| |:|:||||||:
Db           300 AGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTGTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| ||||||||||||::||||||||:|:|||||||||||||||||||||:| |:
Db           360 ACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGTGG 419

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db           420 GAACACCTAT 429


RESULT 805
CQ096442
LOCUS       CQ096442                 433 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 5301 from Patent WO0157272.
ACCESSION   CQ096442
VERSION     CQ096442.1  GI:41065468
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 5301 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL122001.5
                     EXPRESSED IN PLACENTA, SIGNAL = 1.9"

  Query Match             5.1%;  Score 378.6;  DB 1;  Length 433;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  377;  Conservative   38;  Mismatches   17;  Indels    2;  Gaps    1;

Qy       3271 CCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCC 3330
              |||||||:||||:||||||||||:|||||:|||| :|||||||||||||||||||||||
Db             2 CCCAGGTATGGCGAAATAGCCAGGCCATTATATACGCTAATTAAGGAAACTCAGAAAGCC 61

Qy       3331 AATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGTGGCTTTCCAGGCCCCTAAAGAA 3390
              ||||||||||||:|||||||||:|:|| |||:::|||||||||||||||||  |   |
Db            62 AATACCCATTTAGTAAGATGGACACCTAAAGCAGAAGTGGCTTTCCAGGCCCTAACGAAG 121

Qy       3391 GGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAYR 3450
              | |||  ||||||||:|||||| |||||:||||||||:|||||||||||||:||:|||::
Db           122 GCCCT--AACCCAAGCCCCAGTATTAAGCTTGCCAACGGGGCAAGACTTTTCTTTATATG 179

Qy       3451 TCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGGGAYGAGCTTG 3510
              ||||||||||||||||:||:| ||||:|||||||||||||:|||:||| |:|||||||
Db           180 TCACAGAAAAAAACAGGAATAACTCTAGGAGTCCTTACACAGGTCCAAGGAACGAGCTTG 239

Qy       3511 CAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGT 3570
              |||||::||||: ||||||:||||||||:|||||||||||||||||||||:|:|||||||
Db           240 CAACCCGTGGCAGACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTTGGCCTCATTGT 299

Qy       3571 TTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGR 3630
              |||:||||| |||||||||||||| |::|||||||||||||||||||||||||||||| :
Db           300 TTATGGGTAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTTAAAATAATACAGGAA 359

Qy       3631 AGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGAC 3690
              |||||||||||||||||||||||||||||:|||||::|||||||||||||||||||||||
Db           360 AGAGATCTTACTGTGTGGACATCTCATGATGTGAACAGCATACTCACTGCTAAAGGAGAC 419

Qy       3691 TTGTGGCTGTCAGA 3704
              ||||||||||||||
Db           420 CTGTGGCTGTCAGA 433


RESULT 806
CQ747668
LOCUS       CQ747668                 429 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33602 from Patent WO02068579.
ACCESSION   CQ747668
VERSION     CQ747668.1  GI:42371337
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 33602 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.1%;  Score 378.3;  DB 1;  Length 429;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  384;  Conservative   26;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||:||||||||:|||||||||||| ||| |||:|:||| |||||||: |
Db             1 ATGGCATACCTGAGTAAGGAAATTGATGTAGTGGCAGAGGATTGGCCTCACTGTTTACAG 60

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              |||||||  ||||||||||||::|||||||||||||||||||||||||||| ||:||| ||
Db            61 GTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATAATACAAGGAAGAAAT 120

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              |||||||||||||| |||||||||:||||::|||||||||||||||||||||||||||||
```

```
Db        121 CTTACTGTGTGGATATCTCATGATGTGAATGGCATACTCACTGCTAAAGGAGACTTGTGG 180

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
             ||||||||||||:|||||| |||| :|||||||||||||||||||:|||||||||||| :|
Db        181 CTGTCAGACAACCGTTTACCTAAACATCAGGCTCTATTACTTGAAGGGCCAGTGCTGGGA 240

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATA 3816
             ||| | |||||||||||||||||||||:| |||||||||||||||||||||||||||||
Db        241 CTGCGTACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATA 300

Qy       3817 RAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGAR 3876
             :||:|||||||||||||| :|||||| ||||| |||||||||||||||||||||||:||||:
Db        301 GAACATAACTGTCAACAGGTAATTGCTCAAGCCTATGCCACTCGAGGGGACCTTCTAGAG 360

Qy       3877 GTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 3936
             |||||:||||||||||||:|||| |||||||||||||||||||||||||||||||||||||
Db        361 GTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 419

Qy       3937 AGGACTTCGA 3946
             ||||||| ||
Db        420 AGGACTTTGA 429


RESULT 807
CQ068125
LOCUS       CQ068125                422 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 3925 from Patent WO0157278.
ACCESSION   CQ068125
VERSION     CQ068125.1  GI:41037994
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 3925 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN HELA, SIGNAL = 1.1"

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
             ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db          1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
             |  |||||||| ||||||||:||||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
             ||||||||||||||||||||||||||||||||||||:|||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
             ||||||||||||:|||||||||||:|||| |||||||| |||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
             |||||||||||||||||||||||||:||:|||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
             |||||| |||||||||||| ||:||||:||||||:|||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             |||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
             |||
Db        420 TGA 422


RESULT 808
CQ095200
LOCUS       CQ095200                422 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4059 from Patent WO0157272.
ACCESSION   CQ095200
VERSION     CQ095200.1  GI:41064226
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 4059 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN PLACENTA, SIGNAL = 2.6"

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
             ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db          1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
             |  |||||||| ||||||||:||||||||||||||||||||||||||:|||||:: |
Db         61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 3729
             ||||||||||||||||||||||||||||||||||||:|||||||||||:|||||||
Db        121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
             ||||||||||||:|||||||||||:|||| |||||||| |||||||||||||:| ||
Db        181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
             |||||||||||||||||||||||||:||:|||||||| ||||||:||||| ||||||||
Db        241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
             |||||| |||||||||||| ||:||||:||||||:|||||||||||: ||| ||||||||||
Db        301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
             |||||||||||||||||||||||||||||||||||||||||| |||||:||||||||||||||
Db        360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
```

```
                |||
Db        420 TGA 422


RESULT 809
CQ133948
LOCUS       CQ133948                422 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3970 from Patent WO0157276.
ACCESSION   CQ133948
VERSION     CQ133948.1  GI:41091304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 3970 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 0.98"

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |  |||||||| ||||||||:||||||||||||||||||||||||||||:|||||:: |
Db            61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              ||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||
Db           121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              |||||||||||:|||||||||||:|||| |||||||| |||||||||||||:| ||
Db           181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              |||||||||||||||||||||||||:||:|||||||| ||||||:||||| ||||||||
Db           241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGRGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
              |||||| |||||||||| ||:||||:|||||:|||||||||||: ||| ||||||||||
Db           301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
              |||||||||||||||||||||||||||||||||| |||||:||||||||||||||||||
Db           360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
              |||
Db        420 TGA 422


RESULT 810
CQ201644
LOCUS       CQ201644                422 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3921 from Patent WO0157271.
ACCESSION   CQ201644
VERSION     CQ201644.1  GI:41187308
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 3921 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN BT474, SIGNAL = 0.99"

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
              ||||||| |||:|:|||||||||: |||||||||||||||||||::|||| ||||||
Db             1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
              |  |||||||| ||||||||:||||||||||||||||||||||||||||:|||||:: |
Db            61 GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
              ||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||
Db           121 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
              |||||||||||:|||||||||||:|||| |||||||| |||||||||||||:| ||
Db           181 CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
              |||||||||||||||||||||||||:||:|||||||| ||||||:||||| ||||||||
Db           241 TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGRGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
              |||||| |||||||||| ||:||||:|||||:|||||||||||: ||| ||||||||||
Db           301 TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
              |||||||||||||||||||||||||||||||||| |||||:||||||||||||||||||
Db           360 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
              |||
Db        420 TGA 422


RESULT 811
CQ217189
LOCUS       CQ217189                422 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4028 from Patent WO0157273.
ACCESSION   CQ217189
VERSION     CQ217189.1  GI:41198513
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
```

```
          3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
          60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
          August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
          (03.10.00)<150> US 60/236,359<151> 27 September 2000
          (27.09.00)<150> US 60/234,687<151> 21 September 2000
          (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
          Molecular Dynamics Sequence Listing Engine
  JOURNAL Patent: WO 0157273-A 4028 09-AUG-2001;
          Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.4"

  Query Match             5.0%;  Score 377.7;  DB 1;  Length 422;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  384;  Conservative   22;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       3550 GCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAA 3609
Db          1 GCAAAGGATTGGCTTCATTGTTTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTGTCTGAA 60

Qy       3610 GCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGC 3669
Db            GTGGTTAAAATGATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCC 120

Qy       3670 ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCT 3729
Db           ATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCGTTTACTTAAATATCAGGCT 180

Qy       3730 CTATTACTTGAARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCA 3789
Db           CTATTACTTGAAGGGCCAGTGCTGCAACTGCGCACTTGTGTAACTCTTAACCCAGTCACA 240

Qy       3790 TTTCTTCCAGACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACC 3849
Db           TTTCTTCCAGACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACC 300

Qy       3850 TATGCCACTCGAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTA 3909
Db           TATGCCGCTCGAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTA 359

Qy       3910 TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAG 3969
Db           TACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAG 419

Qy       3970 TGA 3972
Db           TGA 422


RESULT 812
AB024484
LOCUS       AB024484                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Papio hamadryas DNA, similar to pol gene of HERV-W and MSRV,
            isolate:HBW6-2.
ACCESSION   AB024484
VERSION     AB024484.1  GI:6063517
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="HBW6-2"
                     /db_xref="taxon:9557"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 377.3;  DB 1;  Length 429;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  383;  Conservative   26;  Mismatches   20;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
Db              AGTGCTGTGACTGCGCACTTGTGCAGCTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
Db              AGAAAAGATAGAATATAACTGTCAACAAGTAATTGCTCAAACCTATGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
Db           CCTTTTAGAGGTTAACTTGACTGACCCCGACC-TCAACTTGTATTCTGGTGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
Db           TTGCAGAAAAAGAACTTCAAAAATCAGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
Db           AAAGTAATCCCCTCACTCCAGGAACTAGTGTTCAGCTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
Db           GGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
Db           ACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTTCTAACTTCCAAGG 419

Qy       4167 GAACACCTAT 4176
Db           GAACACCTAT 429


RESULT 813
AB024467
LOCUS       AB024467                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV,
            isolate:ORW3-1.
ACCESSION   AB024467
VERSION     AB024467.1  GI:6063498
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
```

```
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="ORW3-1"
                     /db_xref="taxon:9600"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 376.1;  DB 1;  Length 429;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  383;  Conservative   25;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  |||| |||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGCGACTGCACACTCGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||:|||||||||||||:||||| |||||||||||||  || |||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||: ||||:|||||||||||:|||| ||||||||||||||||||||||||||
Db        121 CCTTCTAGAGCTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             ||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCGAAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             ||| |||||||||||||||| ||||||| |||: |||:||| |||||||||||| |||:|
Db        240 AAAATAATCCCCTCACTCCAAGAACTAGCGCTCGGCTGGCAAAACTAATAGCCGTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              ||||||| |||||||| |||: |||||||:||||||||||||||||||:|:||||||:
Db        300 AGGCACTAAAATTAGGAAAAGGTAAAAGGGTAAATATATATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| ||||||||||::|||||||:|:|||||||||||||||||||:|||:
Db        360 ACCTAGTCTTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
             ||||||||||
Db        420 GAACACCTAT 429


RESULT 814
AB024471
LOCUS       AB024471                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV,
            isolate:ORW3-5.
ACCESSION   AB024471
VERSION     AB024471.1  GI:6063502
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="ORW3-5"
                     /db_xref="taxon:9600"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 375.7;  DB 1;  Length 429;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  382;  Conservative   26;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||| :||||  |||||||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||:|||||||||||||| :| ||| ||||||||||||||||||||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             || |:|| |:|||||:||||||||||:| || ||||||||||||||||||||||||||||
Db        121 CCGTCTAAAGGTTCCCTTGACTGATCCTG-CCCTCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||| ||||||:|||||||||||||||| ||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGGTTAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||||| |||:||||:||||||||||||:| |:|
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTTATTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             ||||||||||||||||||||:||||||||:||||||||| |||||||||:|:||||||:
Db        300 GGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATAAACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| |||||||||  ::||||||||:|:||||| ||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAAGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
             ||| ||||||
Db        420 GAATACCTAT 429


RESULT 815
AB024474
LOCUS       AB024474                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GIW4-2.
ACCESSION   AB024474
VERSION     AB024474.1  GI:6063505
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-2"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 375.7;  DB 1;  Length 429;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  382;  Conservative   26;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:||||||||||||||||||||||||||||:| ||||||||||||||||||||
Db             1 AGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              |||| |||||:||:|||||||||||  |:||||| |||||||| |  |||||||| |||
Db            61 AGAATAGATAGAACATAACTGTCAATGAGTAATTGCTCAAACCAACTCCACTCGAGAGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              | ||:||||:|||||:||||||||| |: | | ||||||||||||||||||||||||||||
Db           121 CTTTTTAGAGGTTCCCTTGACTGATTCCAATC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||:|||
Db           180 TTGTAGAAAAAGGACTTCGAAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||||||||||||||||||||||||||||:||||:||||||||||||||||:|||:|
Db           240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||||||||| |||||||: |||||||:|||||||||||||||||||:|:|||||||:
Db           300 AGGCACTAGAATTAAGAGAAGGGAAAAGGGTAAATATATATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| |||||||||||::|| |||||:|:|||||||||||||||||||:|| :
Db           360 ACCTAGTCCTCCATGCCCATGCAGCGATATGGAGAGAAAGGGAATTCCTAACTTCTGAAG 419

Qy       4167 GAACACCTAT 4176
              |||||||||
Db           420 AAACACCTAT 429


RESULT 816
CQ751113
LOCUS       CQ751113                429 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 37047 from Patent WO02068579.
ACCESSION   CQ751113
VERSION     CQ751113.1  GI:42383659
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 37047 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.0%;  Score 375.5;  DB 1;  Length 429;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  383;  Conservative   25;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3517 RTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTGRCYTCATTGTTTAYGG 3576
              :||||::|||||| ||||||||:|||||||||||||||||||| :|:||| ||||||: |
Db             1 ATGGCATACCTGATTAAGGAAATTGATGTAGTGGCAAAGGGTTTGCCTCACTGTTTACAG 60

Qy       3577 GTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAATAATACAGGGRAGAGAT 3636
              ||||||| ||||||||||||::|||||||||||||||||||| ||||||||:||||||
Db            61 GTAGTGGCGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAATGATACAGGGAAGAGAT 120

Qy       3637 CTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGAGACTTGTGG 3696
              |||||||||||||||||||||||||:|| ||::|||||||||||||||||||||||||||||
Db           121 CTTACTGTGTGGACATCTCATGATGTCAACGGCATACTCACTGCTAAAGGAGACTTGTGG 180

Qy       3697 CTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGCTGCRA 3756
              ||||||||||||:||||||||||||:||||||||||||||||||||:|||||||| || :
Db           181 CTGTCAGACAACTGTTTACTTAAATATCAGGCTCTATTACTTGAAGGGCCAGTGATGTGG 240

Qy       3757 CTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATA 3816
              |||  |||||||||||||||||||||||:| |||||||||||||||||| ||||||||
Db           241 CTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAAAAAAGATA 300

Qy       3817 RAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGAR 3876
              :||:|||||||||||||| :| ||| ||||||||||||||||||||||||||||:||||:
Db           301 GAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCCACTCGAGGGGACCTTCTAGAG 360

Qy       3877 GTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 3936
              |||||:|||||||||:|||| |||||||||||||||| |||||||||| |||||
Db           361 GTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGAAAGTTCCTTTGTTGAAAA 419

Qy       3937 AGGACTTCGA 3946
              |||| || ||
Db           420 AGGAATTTGA 429


RESULT 817
AB024488
LOCUS       AB024488                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Papio hamadryas DNA, similar to pol gene of HERV-W and MSRV,
            isolate:HBW6-6.
ACCESSION   AB024488
VERSION     AB024488.1  GI:6063521
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
```

```
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="HBW6-6"
                     /db_xref="taxon:9557"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 375.3;  DB 1;  Length 429;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  381;  Conservative   27;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:|||| ||||||||||||||||||||||:| |||||||| || |||||||
Db          1 AGTGCTGCAACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTACAAACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||:|||||||||||||||:||||| |||||||||| ||| | |||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACACCATTAGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:||| |:||||||||||: ||| ||||||| |||||||| |||||||||
Db        121 CCTTCTAGAGGTTTCCTTGACTGATCCCAACC-TCAACTTATATACTGATAGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||| |||||: |||||||||||||||||||| ||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCAAAAAGCAGGGTATGCAGTGGTCAGTGATTATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             ||||||||||||||||||||||||||||||:||||:|||||||| ||||||:|||:|:
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTCATAGCCCTCACTT 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             |||||||||||||||||||||:||||||||:|||||||||||||||||| :|:||||||:
Db        300 GGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATACAGACTCCAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| ||||||||||::||||||||:|:||||||||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
              |||||||||
Db        420 AAACACCTAT 429


RESULT 818
CQ742800
LOCUS       CQ742800                492 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28734 from Patent WO02068579.
ACCESSION   CQ742800
VERSION     CQ742800.1  GI:42354123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 28734 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             5.0%;  Score 374.8;  DB 1;  Length 492;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  396;  Conservative   41;  Mismatches   23;  Indels    4;  Gaps    3;

Qy       3248 AGGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAS 3307
             |||:|||||||||:|||||||| ||||||: |  :|||||||||| :|||||:| |||:
Db         18 AGGTTTCTGCCGAATATGGATT-CCCAGGTACGATGAAATAGCCAGACCATTATACACAC 76

Qy       3308 TAATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAG 3367
             |||||||||||||||||||||||||||||| ||||||:||| |||||:|:|||||:::|||
Db         77 TAATTAAGGAAACTCAGAAAGCCAATACTCATTTAGTAAAATGGACACCTGAAGCAGAAG 136

Qy       3368 TGGCTTTCCAGGCCCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAAC 3427
             |||||||||||||||  ||  | | |  ||||||||:|||||||||||:||||||||
Db        137 TGGCTTTCCAGGCCCTAAAGAAGGCCGT--AACCCAAGCCCCAGTGTTAAGCTTGCCAAC 194

Qy       3428 RGGGCAAGACTTTTSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTA 3487
             :|| ||||||||||:||:|||::|||||| |||||||||:||:|||||:||||||||||
Db        195 AGGACAAGACTTTTCTTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGTCCTTA 253

Qy       3488 CACAGRTCCRAGGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAG 3547
             |||||:|||:|||||:|||||||||| | :||||::|||| |:||||||||:||||||||
Db        254 CACAGATCCGAGGGATGAGCTTGCAATCAGTGGCATACCTAAGTAAGGAAATTGATGTAG 313

Qy       3548 TGGCAAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTG 3607
             |||||||||||||:|:|| |||| ||: |||||||||||||||||||::|||||||||
Db        314 TGGCAAAGGGTTGGCCTCTTTGTCTACAGGTAGTGGTGGCAGTAGCAGTCTTAGTATCTG 373

Qy       3608 AAGCAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYR 3667
             |||||||||||||||||||||:|||||||||||||||||||||||||||||:|||||::
Db        374 AAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATG 433

Qy       3668 GCATACTCACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGT 3711
             ||||| |||||||||||||||||||||||||||||| |||:||
Db        434 GCATATTCACTGCTAAAGGAGACTTGTGGCTGTCAGAGAACTGT 477


RESULT 819
AB024459
LOCUS       AB024459                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pan troglodytes DNA, similar to pol gene of HERV-W and MSRV,
            isolate:CHW1-3.
ACCESSION   AB024459
VERSION     AB024459.1  GI:6063490
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
```

```
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CHW1-3"
                     /db_xref="taxon:9598"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 374.7;  DB 1;  Length 429;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  381;  Conservative   26;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||| :||||||| |||||||||||||||||||:| ||||||||||||||||||||
Db          2 GTGCTGTGACTGTGCGCTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              |||||||||:||:|||||||||||||:||||  |||||||||| ||| || |||| |||
Db         62 GAAAAGATAGAACATAACTGTCAACAAGTAATCGCTCAAACCTACGCCGCTTGAGGAGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||| |||: ||| ||||||||||||||||||||||||||||
Db        122 CTTCTAGAGGTTCCCTTGACTGGTCCCAACC-TCAACTTGTATACTGATGGAAGTTCCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              |||| |||||||||| |||||||:||||||||||||||||| ||||| ||||:||||
Db        181 TGTACAAAAAGGACTTTGAAAAGCGGGGTATGCAGTGGTCAGTGTTAATGAAATACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              |||||||||| |||||||||||||||||:||||:|||||||||||||||:|||:| |
Db        241 AAGTAATCCCCGCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              |||||||||||| |||||||:|||||||:||||||||||||||| ||:|:|:||||||:|
Db        301 GGCACTAGAATTGGGAGAAGGAAAAAGGGTAAATATATATACAGACCCTAAGTATGCTTA 360

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              ||||||| |||||||||| |::||||||||:|:|||||||||||||||||||:|||:|
Db        361 CCTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGG 420

Qy       4168 AACACCTAT 4176
              |||||||||
Db        421 AACACCTAT 429


RESULT 820
AB024480
LOCUS       AB024480                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GIW4-8.
ACCESSION   AB024480
VERSION     AB024480.1  GI:6063513
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-8"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 374.5;  DB 1;  Length 429;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  382;  Conservative   25;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||||||||||||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||: |||||||||||||:||||| |||||||||||||||| |||||||
Db         61 AGAAAAGATAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTATGCCACTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:|||||||||||: ||| ||||||||| ||||| ||||||||| |
Db        121 CCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTGTACTGTTGGAAGTTCTT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| ||||||:|||| |||||||||||||| |||| ||||||:|||
Db        180 TTGTAGAAAAAGGACTTTGAAAAGCGGGGGATGCAGTGGTCAGTAATAAGGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||||||||||||||||||:||||: |||||||||||||:|||:|
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGACAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              ||||||||||||||||||||:| ||||| :|||||||||| ||||||||:|:||||||:
Db        300 GGGCACTAGAATTAGGAGAAGGATAAAGGATAAATATATATGCAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| |||||||||||| :||||||||:|:| ||| ||||||||||||||: ||:
Db        360 ACCTAGTCCTCCATGCCCATGTAGCAATATGGAGAAAAAAGGAATTCCTAACTTCCAAGG 419

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db        420 GAACACCTAT 429


RESULT 821
AB024457
LOCUS       AB024457                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pan troglodytes DNA, similar to pol gene of HERV-W and MSRV,
            isolate:CHW1-1.
ACCESSION   AB024457
VERSION     AB024457.1  GI:6063488
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
```

```
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CHW1-1"
                     /db_xref="taxon:9598"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 374.1;  DB 1;  Length 429;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  381;  Conservative   26;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:|||| ||||||||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGCGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||| | || |:||:|||| |||||||||:||||| |||||||||||||| ||||||||||
Db         61 AGAGAGGAAAGAACATAATTGTCAACAAGTAATTGCTCAAACCTATGCCGCTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|| ||:|||||||||| :|||| ||||||||||||||||||||||||||
Db        121 CCTTTTAGAGGTGCCCTTGACTGATCACGACC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||| |||||: |||| |||||||||||||  ||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCAAAAAGCAGGGTTTGCAGTGGTCAGTCGTAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             ||||||||||||||||||||||||||||||||:||||:||| ||||||||||:||||:|
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAAAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             ||||||||||||||||||||: |||||||:|||||||||||||||||:|:||||||:
Db        300 GGGCACTAGAATTAGGAGAAGGGAAAAGGGTAAATATATATATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| | |||||||||::||||||||:|:|||||||||||||||||| |||:|||:
Db        360 ACCTAGTCCTTCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAATTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
             ||| ||||||
Db        420 GAATACCTAT 429


RESULT 822
AB024460
LOCUS       AB024460                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pan troglodytes DNA, similar to pol gene of HERV-W and MSRV,
            isolate:CHW1-4.
ACCESSION   AB024460
VERSION     AB024460.1  GI:6063491
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CHW1-4"
                     /db_xref="taxon:9598"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 373.1;  DB 1;  Length 429;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  380;  Conservative   26;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:|||| |||||| ||||||||||||||||:| ||||||||||||||||||
Db          2 GTGCTGCAACTGCGCACTTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             |||||||||:||:|||||||||||||:||||| |||||||||  || ||||||||||
Db         62 GAAAAGATAGAATATAACTGTCAACAAGTAATTGCTCAAACCTACCCCGCTCGAGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||| |:|||||| ||| ||||||| |||||||||| |||
Db        122 CTTTTAGAGGTTCCCTTGACTGATTCCGACCTT-AACATGTATACCGATGGAAGTTGCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
             ||||||||||||| | |||||||:|||||||||||||||||| ||||||||||:||||
Db        181 TGTAGAAAAAGGACCTTGAAAAGTGGGGTATGCAGTGGTCAGGGATAATGGAATACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
             || |||||||||||||||||||||||||||||:||||:|||||||||| ||||:|||:| |
Db        241 AAATAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAACAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              |||||||||||||||||:||||| ||:|||||||||| |||||||:|:||||||:|
Db        301 AGCACTAGAATTAGGAGAAGGAAAAATGGTAAATATATATATAGACTCTAAGTATGCTTA 360

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
             ||||||| ||||||||||::||||||||:|:|||||||||||||||||||||||||:|||:|
Db        361 CCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGG 420

Qy       4168 AACACCTAT 4176
              || ||||||
Db        421 AAAACCTAT 429


RESULT 823
AB024470
LOCUS       AB024470                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV,
            isolate:ORW3-4.
ACCESSION   AB024470
VERSION     AB024470.1  GI:6063501
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (sites)
```

```
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="ORW3-4"
                     /db_xref="taxon:9600"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 372.5;  DB 1;  Length 429;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches  380;  Conservative   26;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||| ||:|||||||||||||||||||||||| ||||:| ||||||||||||||||||
Db          1 AGTGCCGCGACTGTGCACTTGTGCAACTCTTAATCCAGTCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||||| |||:||:|||||||||| |||:|||| |||||||  ||| ||||||||||
Db         61 AGAAAAAATAGAACATAACTGTCAGCAAGTAATTGCTCAAACCTGCGCCGCTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||: |||:|||||:|||||||||||: ||| |||||||||||  ||| ||||||||
Db        121 CCTTCCAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACCAATGAAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||||||: |||||||||||||||||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCGAAAAGTTGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||| |||||||||||||||||| |||:||||:|||||||||||||||:|||:|
Db        240 AAAGTAATTCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             ||||||||||||||||||||:||||||||:|||||||||||||||||||:|:||||| :
Db        300 GGGCACTAGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGACTCTAAGTATGCCT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| |||||||||||::|||||||  :|:||| |||||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCATGCAGCAATATAGAGAGACAGGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
             ||||| ||||
Db        420 GAACAGCTAT 429


RESULT 824
AB024487
LOCUS       AB024487                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Papio hamadryas DNA, similar to pol gene of HERV-W and MSRV,
            isolate:HBW6-5.
ACCESSION   AB024487
VERSION     AB024487.1  GI:6063520
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="HBW6-5"
                     /db_xref="taxon:9557"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 372.1;  DB 1;  Length 429;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  379;  Conservative   27;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:|||| |||||||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGCGACTGCGCACTTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||:|||||||||||||:||||| |||||||||||| ||||||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCTACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:|||||||||||: ||| ||| |||||||| ||||||||||||||
Db        121 CCTTCTAGAGGTTCCCTTGACTGATCCCAACC-TCAGCTTGTATATTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             |||||||||||||||||  ||| : |||||||||||||||||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCAGAAAACAGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||||||| |||:||||:|||||||||||||| :|||:|:
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCTCTCACTT 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             |||  |||||| |||||||||:|||| |||:|||||||| ||||||||||:|:|||||:
Db        300 GGGTGCTAGAAATAGGAGAAGGAAAAGGGGTAAATATATGTACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              || |||| |||||| |||||::||||||||:|:|||||||||||||||||||||:|||:
Db        360 GCCCAGTCTTCCATGTCCATGCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
              || |||||||
Db        420 GAGCACCTAT 429


RESULT 825
AB024458
LOCUS       AB024458                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pan troglodytes DNA, similar to pol gene of HERV-W and MSRV,
            isolate:CHW1-2.
ACCESSION   AB024458
VERSION     AB024458.1  GI:6063489
KEYWORDS    .
```

```
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CHW1-2"
                     /db_xref="taxon:9598"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 371.5;  DB 1;  Length 429;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches  379;  Conservative   26;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:||||  |||||| |||||||||||||||||:| ||||||||||||||||||||
Db             2 GTGCTGCAACTGCTCACTTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              |||||||||:||:|||||||||||||||:|||||  ||||||||||   || ||||||||||
Db            62 GAAAAGATAGAACATAACTGTCAACAAGTAATTACTCAAACCTACCCCGCTCGAGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||||| |:|||||| ||| ||||||| |||||||||| |||
Db           122 CTTTTAGAGGTTCCCTTGACTGATTCTGACCTT-AACATGTATACCGATGGAAGTTGCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              |||||||||||||| | ||||||:||||||||||||||||| ||||||||||||:||||
Db           181 TGTAGAAAAAGGACCTTGAAAAGTGGGGTATGCAGTGGTCAGGGATAATGGAATACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              || |||||||||||||||||||||||||||:||||:|||||||||| ||||:|||:| |
Db           241 AAATAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAACAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              ||||||||||||||||||:||||| ||:|||||||||||| |||||||:|:||||||:|
Db           301 AGCACTAGAATTAGGAGAAGGAAAAATGGTAAATATATATATAGACTCTAAGTATGCTTA 360

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              ||||||| ||||||||||||:|||||||:|:|||||||||||||||||||:|||:|
Db           361 CCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGGG 420

Qy       4168 AACACCTAT 4176
              || ||||||
Db           421 AAAACCTAT 429


RESULT 826
AB024462
LOCUS       AB024462                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Gorilla gorilla DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GOW2-1.
ACCESSION   AB024462
VERSION     AB024462.1  GI:6063493
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GOW2-1"
                     /db_xref="taxon:9593"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 370.9;  DB 1;  Length 429;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  379;  Conservative   26;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||||||| |||||||||||||||||||||:| |||||||||||||||| ||
Db             1 AGTGCTGCGACTGTGCGCTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAACGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||: |||||||||||||||:||||| ||||||||||  |  ||||||||||
Db            61 AGAAAAGATAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACACTGCTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||| |||||||:|||| ||||||||||||||||||||||| ||||
Db           121 CCTTCTAGAGGTTCCCTTGGCTGATCCCGACC-TCAACTTGTATACTGATGGAAGCTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||||| ||| |:|| |||||||||||||||||| |||||||||:|||
Db           180 TTGTAGAAAAAGGACTTCAAAAGGCGGAGTATGCAGTGGTCAGTGACAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||||||||||||||||||:|| |:|||||||||||||||:|||:|
Db           240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGGTGGCAGAACTAATAGCCCTCATTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||| |||||||||:||||||| :|||||||||||||| |||:|:||||||:
Db           300 AGGCACTAGAGTTAGGAGAAGGAAAAAGGATAAATATATATACAGAATCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| |||||||| ||:: |||||||:|:|||||||||||||||||||:|||:
Db           360 ACCTAGTCCTCCATGCCCGTGCAACAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGG 419

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db           420 GAACACCTAT 429


RESULT 827
AB024479
```

```
LOCUS       AB024479                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GIW4-7.
ACCESSION   AB024479
VERSION     AB024479.1  GI:6063512
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-7"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             5.0%;  Score 370.9;  DB 1;  Length 429;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches  379;  Conservative   26;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||||:||||  || ||||||||||||||||||:| ||||||||||||  ||||
Db          1 AGTGCTGCGACTGCACATTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGATGATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             ||||||||||:||:||||  ||||||||:||||| ||||||  |||||||||||||||||||
Db         61 AGAAAAGATAGAACATAATGGTCAACAAGTAATTGCTCAAATCTATGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             ||||:||||:|||||:||||||||||: ||| ||| ||||||||||||||||||||| |
Db        121 CCTTCTAGAGGTTCCCTTGACTGATCCCAACC-TCAGCTTGTATACTGATGGAAGTTCTT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCGAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||| |||||||||||||| ||||:| ||: | ||||| ||||||:|||:|
Db        240 AAAGTAATCCCGTCACTCCAGGAACTAATGCTCATCTGACGGAACTGATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             ||||||||||||||||||||:||||||| :||||| ||||||||||||:|:||||||:
Db        300 GGGCACTAGAATTAGGAGAAGGAAAAAGGATAAATACATATACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| ||||||||||:: |||||||:|:||||||||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCATGCAACAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGG 419

Qy       4167 GAACACCTAT 4176
             ||||||||||
Db        420 GAACACCTAT 429
```

```
RESULT 828
AB024461
LOCUS       AB024461                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pan troglodytes DNA, similar to pol gene of HERV-W and MSRV,
            isolate:CHW1-5.
ACCESSION   AB024461
VERSION     AB024461.1  GI:6063492
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CHW1-5"
                     /db_xref="taxon:9598"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.9%;  Score 370.5;  DB 1;  Length 429;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  378;  Conservative   27;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
             ||||||| :||||  ||||||||||||||||||||:| ||||||||||||||||||
Db          1 AGTGCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
             |||| |||||:||:|||||||||||||| :| ||| ||||||||||||||||||||||||
Db         61 AGAAGAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
             || |:||||:|||||:||||||||||:| || ||||||||||||| ||||||||||||||
Db        121 CCATCTAGAGGTTCCCTTGACTGATCCTG-CCCTCAACTTGTATATTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| |||||||:|||||||||||||||||| ||||||||||:|||
Db        180 CTGTAGAAAAAGGACTTTGAAAAGTGGGGTATGCAGTGGTCAGTAATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
             |||||||||||||||||||||||||||||||| ||||:||||:|||||||||||:| |:|:
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTTATTT 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
             |||||||||||||||| ||||:||||||||:||||||||| |||||||||:|:||||||:
Db        300 GGGCACTAGAATTAGGGGAAGGAAAAAGGGTAAATATATAAACAGACTCTAAGTATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
             |||||||| ||||||||||  ::||| ||||:|:||||| ||||||||||||||:|||:
Db        360 ACCTAGTCCTCCATGCCCACACAGCACTATGGAGAGAAAAGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
```

```
                  ||| ||||||
Db        420 GAATACCTAT 429


RESULT 829
AB024481
LOCUS       AB024481                428 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Macaca fuscata DNA, similar to pol gene of HERV-W and MSRV,
            isolate:JMW5-1.
ACCESSION   AB024481
VERSION     AB024481.1  GI:6063514
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 428)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JMW5-1"
                     /db_xref="taxon:9542"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>428
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.9%;  Score 368;  DB 1;  Length 428;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches  383;  Conservative   26;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:|||||||||||||||||||||||||||||:| ||||||||||||||||||||
Db          2 GTGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              |||||||||:||:|||||||||||||:|||| |||||||||| || ||| |||||||
Db         62 GAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACACCGCTCAAGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:|||||||||||:|||| ||||||||||||||||||||||||||||
Db        122 CTTTTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              ||||| |||||||||| ||||||: |||||||||||||||||||||| ||||||:||||
Db        181 TGTAGTAAAAGGACTTTGAAAAGCAGGGTATGCAGTGGTCAGTGATAACGGAATACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              ||||||| |||||||||||||||||||||:||||:| ||||||||||||:|||:|
Db        241 AAGTAAT-CCCTCACTCCAGGAACTAGTGCTCAGCTGGGAGAACTAATAGCCCTCACTCA 299

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              |||||||||||||||||| |:||||||||:||||||||||||||||:|:||||||:|
Db        300 GGCACTAGAATTAGGAGACGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTA 359

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              | ||||| ||||| |||||::||||||||:|:||||||||||||||||||||:|||:|

Db        360 CTTAGTCCTCCATACCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGG 419

Qy       4168 AACACCTAT 4176
              |||||| ||
Db        420 AACACCAAT 428


RESULT 830
AB021921
LOCUS       AB021921                429 bp    DNA     linear   VRL 26-JAN-2000
DEFINITION  Human endogenous retrovirus W gene for polyprotein, partial cds,
            clone:HWX5.
ACCESSION   AB021921
VERSION     AB021921.1  GI:4996595
KEYWORDS    polyprotein.
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (sites)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Identification and phylogeny of novel human endogenous retroviral
            sequences belonging to the HERV-W family on the human X chromosome
  JOURNAL   Arch. Virol. 144 (12), 2403-2413 (1999)
   PUBMED   10664393
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1998) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:87786"
                     /clone="HWX5"
                     /note="HERV-W on human X chromosome"
     CDS             <1. .>429
                     /codon_start=2
                     /product="polyprotein"
                     /protein_id="BAA78549.1"
                     /db_xref="GI:4996596"
                     /translation="LLRLHTCATLNPATFLPDNEEKIEHNCQQVIAQTYAARGDLLEV
                     PLTDPYLNLYTDGSSFVEKGLQKVGYAVVSDNGILESNPHTPGTSTQLAELIALTQAL
                     ELGEQKRVNIYTDSKYVYLVLHAHAAIWRERECLTSKGTP"

  Query Match             4.9%;  Score 367.7;  DB 1;  Length 429;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches  377;  Conservative   26;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              | |||||:|||  |||||||||||||||||||||:| ||||||||||||||||||||
Db          1 ATTGCTGCGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||| :| ||| |||||||||| ||  ||||||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTACGCTGCTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||||: ||| ||||||||||||||||||||||||||
Db        121 CCTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||||| |||||:|||||||||||||||||||||||||||||| :|||
Db        180 TTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATCCTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||| |||||||||||||||  ||:| ||:|||||||||||||:|||:|
Db        240 AAAGTAATCCCCACACTCCAGGAACTAGCACTCAACTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||||||||||||| :|||||||:|||||||||||||||||:|:|||| :
```

```
Db        300 AGGCACTAGAATTAGGAGAACAAAAAAGGGTAAATATATATACAGACTCTAAGTATGTCT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| |||||||||| |::|||||||:|:|||||||||| |||||||||: ||:
Db        360 ACCTAGTCCTCCATGCCCACGCAGCAATATGGAGAGAAAGGGAATGCCTAACTTCCAAGG 419

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db        420 GAACACCTAT 429


RESULT 831
AB024463
LOCUS       AB024463                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Gorilla gorilla DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GOW2-2.
ACCESSION   AB024463
VERSION     AB024463.1  GI:6063494
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GOW2-2"
                     /db_xref="taxon:9593"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.9%;  Score 366.7;  DB 1;  Length 429;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  376;  Conservative   26;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:| || |||||| |||||||||||||||||:| |||||||||||||||||||
Db          2 GTGCTGCAATTGCGCACTTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              ||||| |||:||:||||||| ||||||:||||| |||| |||||  || ||||||||||
Db         62 GAAAATATAGAACATAACTGCCAACAAGTAATTGCTCACACCTACCCCGCTCGAGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||||| |:|||||| |||||||||| | ||||||||||||
Db        122 CTTTTAGAGGTTCCCTTGACTGATTCCGACCTT-AACTTGTATACCGGTGGAAGTTCCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              |||||||||||||| | ||||||:|||||||||| |||||| ||||||||||||:||||
Db        181 TGTAGAAAAAGGACCTTGAAAAGTGGGGTATGCAGCGGTCAGGGATAATGGAATACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              |||||||||||||||| ||||||||||||:||||:||||||||||||||:|||:| ||
Db        241 AAGTAATCCCCTCACTCCGGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              |||||||||||||||| |||:||||||||:|||||||||||||||||||:|:||||||:|
Db        301 GGCACTAGAATTAGGAAAAGGAAAAAGGGTAAATATATATACAGACTCTAAGTATGCTTA 360

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              ||||||| ||||||||||||::||||||||:|:|||| |||||||||||| |||:|||:|
Db        361 CCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAGGGGAATTCCTAATTTCTGAGGG 420

Qy       4168 AACACCTAT 4176
              |  ||||||
Db        421 AGAACCTAT 429


RESULT 832
AB024464
LOCUS       AB024464                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Gorilla gorilla DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GOW2-3.
ACCESSION   AB024464
VERSION     AB024464.1  GI:6063495
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GOW2-3"
                     /db_xref="taxon:9593"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.9%;  Score 366.7;  DB 1;  Length 429;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  376;  Conservative   26;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:| || |||||| |||||||||||||||||:| |||||||||||||||||||
Db          2 GTGCTGCAATTGCGCACTTTTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              ||||| |||:||:||||||| ||||||:||||| |||| |||||  || ||||||||||
Db         62 GAAAATATAGAACATAACTGCCAACAAGTAATTGCTCACACCTACCCCGCTCGAGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||||| |:|||||| |||||||||| | ||||||||||||
Db        122 CTTTTAGAGGTTCCCTTGACTGATTCCGACCTT-AACTTGTATACCGGTGGAAGTTCCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              |||||||||||||| | ||||||:|||||||||| |||||| ||||||||||||:|||
Db        181 TGTAGAAAAAGGACCTTGAAAAGTGGGGTATGCAGCGGTCAGGGATAATGGAATACTTGG 240
```

```
Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              || |||||||||||||||||||| ||||||:||||:||||||||||||||||:|||:| |
Db        241 AAATAATCCCCTCACTCCAGGAACCAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              ||||||||||||||||||:|||||||:|||||||||| |||||||:|:||||||:|
Db        301 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATATAGACTCTAAGTATGCTTA 360

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              ||||||| |||||||||||::||||||||:|:||||||||||||||||||:|||:|
Db        361 CCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGG 420

Qy       4168 AACACCTAT 4176
              |  ||||||
Db        421 AGAACCTAT 429
```

```
RESULT 833
AB024476
LOCUS       AB024476                428 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GIW4-4.
ACCESSION   AB024476
VERSION     AB024476.1  GI:6063507
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 428)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-4"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>428
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.9%;  Score 363.8;  DB 1;  Length 428;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches  380;  Conservative   27;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              | ||||||:||||  |||||||||||||||||||||:| ||||||||||||||||||
Db          1 AATGCTGCGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||| :| ||| |||||| ||| ||||||||||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAAACTACGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||:|||| ||||||||||||||||||||| ||||
Db        121 CCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGCTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              | |||||||||||||||| |||||:|||||||||||||||||||||||||||||:|||
Db        180 TGGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||| || ||||||||||:|||||:|||| |||||:|||:|:
Db        240 AAAGTAATCCCCTCACTCTAGAAACTAGTGCTCAGCTGGCAGAACTAGTAGCCCTCACTT 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||| ||||·|||||||:|||||||:|| ||||||||||||| ||:|:||| ||:
Db        300 GGGCACTAAAATT-GGAGAAGAAAAAAGGGTAAGTATATATACAGACACTAAGTATACTT 358

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              | |||||| |||||||||||::||||||||:|:||||||||||||||||||||:|||:
Db        359 ATCTAGTCATCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCTGAGG 418

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db        419 GAACACCTAT 428
```

```
RESULT 834
AB024465
LOCUS       AB024465                426 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Gorilla gorilla DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GOW2-4.
ACCESSION   AB024465
VERSION     AB024465.1  GI:6063496
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 426)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GOW2-4"
                     /db_xref="taxon:9593"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>426
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.8%;  Score 362.2;  DB 1;  Length 426;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  380;  Conservative   25;  Mismatches   20;  Indels    4;  Gaps    2;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||||:|||| |||||||||||||||||||||:| |||||||||||||| |||
Db          2 GTGCTGCGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAAAGAA 61

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
              |||||||||:||:|||||||||||||||:||||| ||||||||||||  |||  ||||||
Db         62 GAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTATGCTGCTCAGGGGGAC 121

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
              |||:||||:|||||:||||||||||: ||| |||||||||||||| ||||||||||||
```

```
Db        122 CTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTAATGGAAGTTCCTT 180

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
              ||||||||||||||||||||||||:||||||||||||| ||||||||||||||| ||:||||
Db        181 TGTAGAAAAAGGACTTCGAAAAGCGGGGTATGCAGTTGTCAGTGATAATGGAGTACTTGA 240

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
              ||||||||||||||||||||||||||||||||:||||:| | |||||||||||:|||:| |
Db        241 AAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGTAAAACTAATAGCCCTCACTCG 300

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYA 4107
              |||||||||||||  ||||||:||||||||: |||||||||||||| |    |:||||||:|
Db        301 GGCACTAGAATTAAGAGAAGGAAAAAGGGTGAATATATATACAGAGT---AGTATGCTTA 357

Qy       4108 CCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRG 4167
              ||| ||| ||||||||||||::||||||||:|:||||||||||||||||||||:|||:|
Db        358 CCTGGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTCCTAACTTCCGAGGG 417

Qy       4168 AACACCTAT 4176
              |||||||||
Db        418 AACACCTAT 426


RESULT 835
AB024466
LOCUS       AB024466               431 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Gorilla gorilla DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GOW2-5.
ACCESSION   AB024466
VERSION     AB024466.1  GI:6063497
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 431)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GOW2-5"
                     /db_xref="taxon:9593"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>431
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.8%;  Score 361.5;  DB 1;  Length 431;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  380;  Conservative   26;  Mismatches   23;  Indels    3;  Gaps    2;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||| |||||||||||||||||||||||:| |||||||| ||||||||||
Db          1 AGTGCTGCGACTGCGCACTTGTGCAACTCTTAACCCAGCCACATTTCTTTCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              || |||||||:||:|||||||||||||| :| ||| |||||||||| ||||||||||||||
Db         61 AGGAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTACGCCACTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||: ||| ||||||||| |||||||||||| ||||
Db         121 CCTTCTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTGTACTGATGGAAGCTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| |||| |:|||||| |||||||||||||||||||||||:|||
Db         180 TTGTAGAAAAAGGACTTTGAAAGGTGGGGTACGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||  |||||||||||||||| |||:||||:|||||||||||||||:|||:|
Db         240 AAAGTAATCCCTTCACTCCAGGAACTAGCGCTCAGCTGGCAGAACTAATAGCCCTCACTC 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGACTCTRARTATGC 4104
              |||||||||||||||||||:||||||||:||  ||||||||||||||||:|:|||||
Db         300 AGGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATATACAGACTCTAAGTATGC 359

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
              |:||| |||| |||||||||| |::|||||||:|:||| ||| ||||||||||||:||
Db         360 TTACCCAGTCCTCCATGCCCACGCAGCAATATGGAGAGACAGGAAATTCCTAACTTCTGA 419

Qy       4165 GRGAACACCTAT 4176
              |:||||||||||
Db         420 GGGAACACCTAT 431


RESULT 836
AB024486
LOCUS       AB024486               427 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Papio hamadryas DNA, similar to pol gene of HERV-W and MSRV,
            isolate:HBW6-4.
ACCESSION   AB024486
VERSION     AB024486.1  GI:6063519
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 427)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="HBW6-4"
                     /db_xref="taxon:9557"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>427
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.8%;  Score 359.5;  DB 1;  Length 427;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  385;  Conservative   25;  Mismatches   17;  Indels    3;  Gaps    3;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:||| ||||||||||||||||||||||||| |||||:| ||||||||||||||
Db          1 AGTGCTGCGACTACGCACTTGTGCAACTCTTAA-CCAGCCACATTTCTTCCAGACAATGA 59
```

```
Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||||:||||| |||||||||| ||| ||||||||||
Db         60 AGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAACCTACGCCGCTCGAGGGGA 119

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:|||||||||||:|||| ||||||||||||||||||||||||||
Db        120 CCTTTTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAAGTTCCT 178

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              ||||||||||||||||||||||:|||||||| |||||||||| |||||||||:|||
Db        179 TTGTAGAAAAAGGACTTCGAAAAGCGGGGTATGTAGTGGTCAGTGAAAATGGAATACTTG 238

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||||||||||| ||||||||||||:||||:|||||||||||||:|||:|
Db        239 AAAGTAATCCCCTCACTTCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 298

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              ||||| |||| ||||||||||:||||| ||:|||||||||||||||||:|:||||||:
Db        299 GGGCA-TAGAGTTAGGAGAAGAAAAAAAGGTAAATATATATACAGACTCTAAGTATGCTT 357

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||| | |||||| ||||| :||||||||:|:|| |||||||||||||||||:|||:
Db        358 ACCTAGCCCTCCATGTCCATGTAGCAATATGGAAAGGAAGGGAATTCCTAACTTCCGAGG 417

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db        418 GAACACCTAT 427
```

RESULT 837
AB024475 .

```
LOCUS       AB024475                429 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV,
            isolate:GIW4-3.
ACCESSION   AB024475
VERSION     AB024475.1  GI:6063506
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 429)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:4418652244990)
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GIW4-3"
                     /db_xref="taxon:9579"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>429
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.8%;  Score 357.7;  DB 1;  Length 429;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  370;  Conservative   27;  Mismatches   32;  Indels    1;  Gaps    1;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||| :|||| || |||||||| ||||||||||||:| |||||||| ||||||||||
Db          1 AGTGCTGTGACTGCGCTCTTGTGCAGCTCTTAACCCAGCCACATTTCTTTCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||||:||||| |||||||||| ||| ||||||||||
Db         61 AGAAAAGATAGAACATAACTGTTAACAAGTAATTGCTCAAACCTATGCCGCTAGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:|||||||||||:|||| |||||| |||||||||| ||||| | |
Db        121 CCTTTTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTCGTATACTGATTGAAGTCCTT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              ||      ||| | ||||| ||||||:| ||||||||||||||||||||||||||:|||
Db        180 TTTCTACAAAGGAACTTCTAAAAGCGAGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||||||||||||||||||||||||||||:| ||: |||||||||||||:|||:|:
Db        240 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTCACCTGTCAGAACTAATAGCCCTCACTT 299

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||||||||||||||||||: |||||| :|||||||||||||||| |:|: |||||:
Db        300 GGGCACTAGAATTAGGAGAAGGGAAAAGGATAAATATATATACAGACTTTAAGAATGCTT 359

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              ||||  || ||||||||||||::|||||||:|:|||||||||||||||||||||:|||:
Db        360 ACCTCATCCTCCATGCCCATGCAGCAATATGGAAAGAAAGGGAATTCCTAACTTCCGAGG 419

Qy       4167 GAACACCTAT 4176
              ||||||||||
Db        420 GAACACCTAT 429
```

RESULT 838
AB055868

```
LOCUS       AB055868                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:2F7-2.
ACCESSION   AB055868
VERSION     AB055868.1  GI:15207777
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="2F7-2"
                     /db_xref="taxon:9606"
                     /cell_line="2F7"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60
```

```
Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| ||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAAGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

RESULT 839
AB055876
LOCUS AB055876 373 bp DNA linear PRI 16-AUG-2001
DEFINITION Homo sapiens DNA, HERV-W LTR, isolate:SiHa-8.
ACCESSION AB055876
VERSION AB055876.1 GI:15207785
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lee,J. and Kim,H.
TITLE Identification and phylogeny of new human endogenous retrovirus HERV-W family in cancer cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 373)
AUTHORS Lee,J. and Kim,H.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346, Fax:82515812962)
FEATURES Location/Qualifiers
source 1. .373
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="SiHa-8"
/db_xref="taxon:9606"
/cell_line="SiHa"
LTR 1. .373
/note="HERV-W family"

Query Match 4.8%; Score 357.5; DB 1; Length 373;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 365; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

```
Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| ||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAAGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

RESULT 840
AB064529
LOCUS AB064529 373 bp DNA linear PRI 05-JAN-2002
DEFINITION Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X3-1.
ACCESSION AB064529
VERSION AB064529.1 GI:18148361
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kim,H. and Crow,T.J.
TITLE Isolation of HERV-W LTR in Xq21.3 region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 373)
AUTHORS Kim,H. and Crow,T.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University, Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346, Fax:82-51-581-2962)
FEATURES Location/Qualifiers
source 1. .373
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="X3-1"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq21.3"
LTR 1. .373
/note="HERV-W LTR"

Query Match 4.8%; Score 357.5; DB 1; Length 373;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 365; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

```
Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240
```

```
Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 841
AB064533
LOCUS       AB064533                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X7-2.
ACCESSION   AB064533
VERSION     AB064533.1  GI:18148365
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X7-2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 842
AB064538
LOCUS       AB064538                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X11-2.
ACCESSION   AB064538
VERSION     AB064538.1  GI:18148370
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X11-2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

```
RESULT 843
AB064539
LOCUS       AB064539                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X13-1.
ACCESSION   AB064539
VERSION     AB064539.1  GI:18148371
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X13-1"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 844
AB066648
LOCUS       AB066648                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HB-1.
ACCESSION   AB066648
VERSION     AB066648.1  GI:18181890
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-1"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 845
AB066659
LOCUS       AB066659                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-12.
ACCESSION   AB066659
VERSION     AB066659.1  GI:18181901
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-12"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 846
AB071972
LOCUS       AB071972                373 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch3.
ACCESSION   AB071972
VERSION     AB071972.1  GI:19698207
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch3"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 357.5;  DB 1;  Length 373;
  Best Local Similarity   97.6%;  Pred. No. 0;
  Matches  365;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 847
AB066205
LOCUS       AB066205                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR, isolate:pla-7.
ACCESSION   AB066205
VERSION     AB066205.1  GI:18181889
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.-S.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="pla-7"
```

```
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 356.5;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||| |||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTNCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 848
AB055871
LOCUS       AB055871                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:A549-3.
ACCESSION   AB055871
VERSION     AB055871.1  GI:15207780
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="A549-3"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||| |||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTCGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 849
AB055872
LOCUS       AB055872                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:HepG2-3.
ACCESSION   AB055872
VERSION     AB055872.1  GI:15207781
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="HepG2-3"
                     /db_xref="taxon:9606"
                     /cell_line="HepG2"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||| |||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60
```

```
Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTTTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| ||||||||||||||||||||:||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

```
RESULT 850
AB055874
LOCUS       AB055874                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:PC3-1.
ACCESSION   AB055874
VERSION     AB055874.1  GI:15207783
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="PC3-1"
                     /db_xref="taxon:9606"
                     /cell_line="PC3"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180
```

```
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| ||||||||||||||||||||:||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

```
RESULT 851
AB064530
LOCUS       AB064530                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X3-2.
ACCESSION   AB064530
VERSION     AB064530.1  GI:18148362
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X3-2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACTTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240
```

```
Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 852
AB064531
LOCUS       AB064531                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X3-3.
ACCESSION   AB064531
VERSION     AB064531.1  GI:18148363
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X3-3"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||| ||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCGTCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 853
AB064535
LOCUS       AB064535                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X7-4.
ACCESSION   AB064535
VERSION     AB064535.1  GI:18148367
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X7-4"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||| ||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCTACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373
```

```
RESULT 854
AB064537
LOCUS       AB064537                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X11-1.
ACCESSION   AB064537
VERSION     AB064537.1  GI:18148369
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X11-1"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGTAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 855
AB066204
LOCUS       AB066204                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR, isolate:pla-6.
ACCESSION   AB066204
VERSION     AB066204.1  GI:18181888
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.-S.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="pla-6"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 TGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 856
AB066655
LOCUS       AB066655                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-4.
ACCESSION   AB066655
VERSION     AB066655.1  GI:18181897
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-4"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGTTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 857
AB066656
LOCUS       AB066656                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-8.
ACCESSION   AB066656
VERSION     AB066656.1  GI:18181898
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-8"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGTGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 858
AB066657
LOCUS       AB066657                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-9.
ACCESSION   AB066657
VERSION     AB066657.1  GI:18181899
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-9"
                     /db_xref="taxon:9606"
```

```
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.8%;  Score 355.9;  DB 1;  Length 373;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||| |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGATGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 859
AB055873
LOCUS       AB055873                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:MP2-2.
ACCESSION   AB055873
VERSION     AB055873.1  GI:15207782
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="MP2-2"
                     /db_xref="taxon:9606"
                     /cell_line="MIA-PaCa-2"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTTTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||:|||||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACCCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 860
AB055877
LOCUS       AB055877                373 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:SiHa-10.
ACCESSION   AB055877
VERSION     AB055877.1  GI:15207786
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="SiHa-10"
                     /db_xref="taxon:9606"
                     /cell_line="SiHa"
     LTR             1. .373
                     /note="HERV-W family"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy         268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
                |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy         328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
                |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy         388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
                |||||||||||||:|||||||||||||||||||||||||||||||||| ||||||||||||||
Db         241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTTCATTCCTTGGAATC 300

Qy         448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
                |||||||:|||  ||||||||||||||||||||||:|||||||||||||||||||||||||
Db         301 CGTGAGGCCAA-AAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy         508 CTGCTACCRTCTTG 521
                |||||||:|||||
Db         360 CTGCTACCATCTTG 373


RESULT 861
AB064536
LOCUS       AB064536                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X7-5.
ACCESSION   AB064536
VERSION     AB064536.1  GI:18148368
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X7-5"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy         148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db           1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy         208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
                |||||||||||||||||||||||| |||||||||||||||||||||||||||||||:||||
Db          61 TGACTCCCATCCCTCTGGATCCTGCGGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy         268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         121 TCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy         328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
                |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy         388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
                |||||||||||||:|||||||||||||||||||||||||||||||||| ||||||||||||||
Db         241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy         448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
                |||||||:|||  ||||||||||||||||||||||:|||||||||||||||||||||||||
Db         301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy         508 CTGCTACCRTCTTG 521
                |||||||:|||||
Db         360 CTGCTACCATCTTG 373


RESULT 862
AB064541
LOCUS       AB064541                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X13-3.
            Xq21.3 region.
ACCESSION   AB064541
VERSION     AB064541.1  GI:18148373
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X13-3"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy         148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db           1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy         208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db          61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy         268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCCGCACGGCTAAGTGCCTG 327
                |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db         121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCCGCACGGCTAAGTGCCTG 180

Qy         328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
                |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240
```

```
Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||| |||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCTCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:|||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 863
AB064542
LOCUS       AB064542                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X13-4.
ACCESSION   AB064542
VERSION     AB064542.1  GI:18148374
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X13-4"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGGGTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:|||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 864
AB066653
LOCUS       AB066653                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-1.
ACCESSION   AB066653
VERSION     AB066653.1  GI:18181895
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-1"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:|||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:| | |
Db        360 CTGCTACCATTTGG 373


RESULT 865
```

```
AB066654
LOCUS       AB066654                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-2.
ACCESSION   AB066654
VERSION     AB066654.1  GI:18181896
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-2"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.7%;  Score 354.3;  DB 1;  Length 373;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||| ||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGATGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGGCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 866
AB066203
LOCUS       AB066203                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR, isolate:pla-5.
ACCESSION   AB066203
VERSION     AB066203.1  GI:18181887
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.-S.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="pla-5"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 353.3;  DB 1;  Length 373;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  362;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||| |||||||||||||||||||| || ||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCNCTGCTGTTTGCCACCACCGCAGACTTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 867
BD058454
LOCUS       BD058454                384 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Secreted expressed sequence tags (sESTs).
ACCESSION   BD058454
VERSION     BD058454.1  GI:22604060
KEYWORDS    JP 2001519666-A/309.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Spaulding,V. and Agostino,M.J.
  TITLE     Secreted expressed sequence tags (sESTs)
  JOURNAL   Patent: JP 2001519666-A 309 23-OCT-2001;
```

```
            GENETICS INSTITUTE INC
COMMENT     PN   JP 2001519666-A/309
            PD   23-OCT-2001
            PF   10-APR-1998 JP 1998543068
            PR   10-APR-1997 US    08/835913
            PI   KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
              DAVID MERBERG,
            PI   MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
            C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC   Strandedness:
            Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .384
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"

  Query Match             4.7%;  Score 352.9;  DB 1;  Length 384;
  Best Local Similarity   96.0%;  Pred. No. 0;
  Matches  361;  Conservative    9;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 5471
              ||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCTGCCTCGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAA 60

Qy       5472 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 5531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTG 120

Qy       5532 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAATCCCCCATGGSCCTCCC 5591
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db        121 CTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT-CCCCATGGCCCTCCC 179

Qy       5592 TTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACTGCACCCCCTCC 5651
              ||||||||||||||||:||:||||:|||:||||||:||||||||||||||||||||||||
Db        180 TTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCCCCTCC 239

Qy       5652 ATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGAATGCAGCGTCC 5711
              |||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||||
Db        240 ATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCC 299

Qy       5712 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACCTTCACTGC 5771
              |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        300 CGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC 359

Qy       5772 CCACACCCATATGCCC 5787
              ||||||||||||||||
Db        360 CCACACCCATATGCCC 375


RESULT 868
AB064540
LOCUS       AB064540                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X13-2.
ACCESSION   AB064540
VERSION     AB064540.1  GI:18148372
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES              Location/Qualifiers
     source           1. .373
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /isolate="X13-2"
                      /db_xref="taxon:9606"
                      /chromosome="X"
                      /map="Xq21.3"
     LTR              1. .373
                      /note="HERV-W LTR"

  Query Match             4.7%;  Score 352.7;  DB 1;  Length 373;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  362;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCCCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAACTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||||:||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 869
AB066201
LOCUS       AB066201                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR, isolate:pla-1.
ACCESSION   AB066201
VERSION     AB066201.1  GI:18181885
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.-S.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES              Location/Qualifiers
     source           1. .373
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /isolate="pla-1"
```

```
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 352.7;  DB 1;  Length 373;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches  362;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||| |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTCAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||||||||||:|||||||||||||||||||||||||||||| ||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTTCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:|||  |||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-AAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 870
AB024468
LOCUS       AB024468                427 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Pongo pygmaeus DNA, similar to pol gene of HERV-W and MSRV,
            isolate:ORW3-2.
ACCESSION   AB024468
VERSION     AB024468.1  GI:6063499
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 427)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="ORW3-2"
                     /db_xref="taxon:9600"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>427
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.7%;  Score 351.5;  DB 1;  Length 427;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  373;  Conservative   26;  Mismatches   28;  Indels    3;  Gaps    2;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:||||  |||||||||||||||  || || |:| ||||||||||||||||||
Db          1 AGTGCTGCGACTGCACACTTGTGCAACTC--AATCCCGCCACATTTCTTCCAGACAATGA 58

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              |||||| |||:||:|||||||||||||:|| || |||||||||  || ||| ||||||
Db         59 AGAAAAAATAGAACATAACTGTCAACAAGTAGTTGCTCAAACCTACACCGCTCAAGGGGA 118

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||:|||| ||||||||||||||||||||||||||
Db        119 CCTTCTAGAGGTTCCCTTGACTGATCCTGACC-TCAACTTGTATACTGATGGAAGTTCCT 177

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||| ||||||: |||||||||||||||||||||||||||:|||
Db        178 TTGTAGAAAAAGGACTTTGAAAAGCAGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 237

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||| |||||||||||||||||||||||:||||:||||||||||||||:|||:|
Db        238 AAAGCAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTC 297

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              |||| || |||||||||||||:||||||||:|||||||||||||||||| |: :| ||||:
Db        298 GGGCCCTCGAATTAGGAGAAGAAAAAAGGGTAAATATATATACAGACTGTATGTGTGCTT 357

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              || ||||| |||||||||| ::|||||||:|:|||||||||||||||||||| |: ||:
Db        358 ACTTAGTCCTCCATGCCCACACAGCAATATGGAGAGAAAGGGAATTCCTAACTCCCAAGG 417

Qy       4167 GAACACCTAT 4176
              |||||||||
Db        418 AAACACCTAT 427


RESULT 871
AB064534
LOCUS       AB064534                373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X7-3.
ACCESSION   AB064534
VERSION     AB064534.1  GI:18148366
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X7-3"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"
```

```
  Query Match             4.7%;  Score 351.1;  DB 1;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  361;  Conservative    5;  Mismatches    7;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGTTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||| |||||||||||||||||||||||||||||| ||||||||:||||
Db         61 TGACTCCCATCCCTTTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGGTCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||||| |||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCCTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 872
AB066202
LOCUS       AB066202                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR, isolate:pla-4.
ACCESSION   AB066202
VERSION     AB066202.1  GI:18181886
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.-S.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished .
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
            Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="pla-4"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.7%;  Score 351.1;  DB 1;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches  361;  Conservative    5;  Mismatches    7;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||  |||||||||||||||||||||||||||||||||| ||| |||||
Db          1 CGAGCTGAGCTTACGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCAGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||||
Db        301 TGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 873
AB024483
LOCUS       AB024483                427 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Papio hamadryas DNA, similar to pol gene of HERV-W and MSRV,
            isolate:HBW6-1.
ACCESSION   AB024483
VERSION     AB024483.1  GI:6063516
KEYWORDS    .
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 427)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="HBW6-1"
                     /db_xref="taxon:9557"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>427
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.7%;  Score 350.9;  DB 1;  Length 427;
  Best Local Similarity   87.9%;  Pred. No. 0;
  Matches  379;  Conservative   26;  Mismatches   19;  Indels    7;  Gaps    3;

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
              |||||| :|||| |||||||||||||||||:| |||||||||||||||||||
Db          2 GTGCTGTGACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGAA 61
```

```
Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
Db         62 GAAAAGATAGAACATAATTGTCAAC----AGTTGCTCAAACCTATGCCACTCGAAGGGAC 117

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
Db        118 CTTTTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGAGGGAAGCTCCTT 176

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGA 3987
Db        177 TGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGAATACTTGA 236

Qy       3988 AAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKG 4047
Db        237 AGGTAATCCCCTCACTCTAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCCCTCACTTG 296

Qy       4048 GGCACTAGAATTAGGAGAAGRAAAAAGGGYAA--ATATATATACAGACTCTRARTATGCT 4105
Db        297 GGCACTAGAATTAGGAGAAGGAAAAAGGGTAAATATATATATACAGACTTTAAGTATGCT 356

Qy       4106 YACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAG 4165
Db        357 TACCTAATCCTCCATGCCCACACAGCAATATGGAGAGAAAGAGAATTCCTAACTTCCGAG 416

Qy       4166 RGAACACCTAT 4176
Db        417 GGAACACCTAT 427
```

```
RESULT 874
AB079272
LOCUS       AB079272                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-4.
ACCESSION   AB079272
VERSION     AB079272.1  GI:28201103
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-4"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             4.7%;  Score 349.5;  DB 1;  Length 373;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  360;  Conservative    5;  Mismatches    8;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
Db         61 CGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCCG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
Db        181 GGTTTGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
Db        241 ACGGCTTCTAATAGAACTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
Db        360 CTGCTACCATCTTG 373
```

```
RESULT 875
AB066649
LOCUS       AB066649                373 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HB-3.
ACCESSION   AB066649
VERSION     AB066649.1  GI:18181891
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-3"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR family"

  Query Match             4.6%;  Score 347.9;  DB 1;  Length 373;
  Best Local Similarity   96.0%;  Pred. No. 0;
  Matches  359;  Conservative    5;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCCGTGACCC 240
```

```
Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| |||||||||| |||||||||||||| ||||||||||||||||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGGTTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              | |||||:||| |||||||||||||||||| ||||:|||||||||||||||||||||||
Db        301 CATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              ||||||||:|||||
Db        360 CTGCTACCATCTTG 373


RESULT 876
AB071975
LOCUS       AB071975                374 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch6.
ACCESSION   AB071975
VERSION     AB071975.1  GI:19698210
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 374)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch6"
                     /db_xref="taxon:9606"
     LTR             1. .374
                     /note="HERV-W LTR"

  Query Match             4.6%;  Score 345.4;  DB 1;  Length 374;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCA-TTGTNCCTGCACGGCTAAGTGCCT 326
              |||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTTGTTCCTGCACGGCTAAGTGCCT 180

Qy        327 GGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACC 386
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACC 240

Qy        387 CACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 446
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 300

Qy        447 CCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGG 506
              ||||||||:||| |||||||||||||||||||||||:||||||||||||||||||||||
Db        301 CCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGG 359

Qy        507 CCTGCTACCRTCTTG 521
              |||||||||:||| |
Db        360 CCTGCTACCATCTGG 374


RESULT 877
AR344412
LOCUS       AR344412                429 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 138 from patent US 6582703.
ACCESSION   AR344412
VERSION     AR344412.1  GI:33740353
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 138 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             4.6%;  Score 345;  DB 1;  Length 429;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  361;  Conservative   41;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| | ||:|||:| |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCCTCAGTATGGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| || |||||||||||||:|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCACCAAGCCACCCAAGCGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |:|:|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:|||||||| |||||||:::|||||||:|
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||||||||| ||||:|||||||||||||||||||:::::|||||||||:|||:|
Db        232 ATACTGGCTTATCCCCATCCCAAAACCCTAAAGCAACTAAGARGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||||| |: ·||:||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCAGATACAGCGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |  |||||||||||||||||||||||||||||||||||||||||||||||||:|||||:|||||| :::||||
Db        351 ATCTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 878
AR344409
LOCUS       AR344409                429 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 135 from patent US 6582703.
ACCESSION   AR344409
VERSION     AR344409.1  GI:33740350
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 135 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             4.6%;  Score 344.6;  DB 1;  Length 429;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  347;  Conservative   60;  Mismatches   21;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| |:||:|||:: ||||||:||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCYTCAGTATRGGGATGAYTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
               |:||: ||::|||||||||||||| ||:|||||||||||::|||||:||||||||:
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCAYCAAGCCACCCAAGYGCTCTTAAATTTCCTY 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||         ||||||| |:::|||||:||||||||||:|||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAARGGCTCASCTCTGCTCACASCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||| |:|||:|||||||| ||||||:::||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCRCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              |||||||:||||||:||||:||:||||:|||||||||||||::::|||||||||:||::|
Db        232 ATACTGGYTTATCCYCATCCCAWAACCYTAAAGCAACTAAGARGGTTCCTTGGCATAWCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||: | |:  |::|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCV-GATACAGYGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |::|||||||||||||||||||||||||| ||:|||||||||:|:|||||: :::|||||
Db        351 ADYTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGARACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              ||||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 879
AB024482
LOCUS       AB024482                428 bp    DNA     linear   PRI 21-OCT-1999
DEFINITION  Macaca fuscata DNA, similar to pol gene of HERV-W and MSRV,
            isolate:JMW5-2.
ACCESSION   AB024482
VERSION     AB024482.1  GI:6063515
KEYWORDS    .
SOURCE      Macaca fuscata (Japanese macaque)
  ORGANISM  Macaca fuscata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Kim,H.S., Takenaka,O. and Crow,T.J.
  TITLE     Isolation and phylogeny of endogenous retrovirus sequences
            belonging to the HERV-W family in primates
  JOURNAL   J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
   PUBMED   10573154
REFERENCE   2  (bases 1 to 428)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Macaca fuscata"
                     /mol_type="genomic DNA"
                     /isolate="JMW5-2"
                     /db_xref="taxon:9542"
                     /sex="Male"
                     /tissue_type="Blood"
     misc_feature    <1. .>428
                     /note="similar to pol gene of human endogenous retrovirus
                     family (HERV-W) and multiple sclerosis-associated
                     retrovirus (MSRV)"

  Query Match             4.6%;  Score 343;  DB 1;  Length 428;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  380;  Conservative   27;  Mismatches   19;  Indels    6;  Gaps    4;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              |||||| |:|||| |||||| |||||||||| ||:| |||||| ||||||||||||
Db          1 AGTGCTACGACTGCATACTTGTGTAACTCTTAACCTAGCCACATTTCCTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||:||||| |||||| ||||||| ||||||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAAGTAATTGCTCAAATCTATGCCGCTCGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:||| |:|||||||||||: ||| ||||||| ||||||||||||||||||
Db        121 CCTTCTAGAGGTTTCCTTGACTGATCCCAACC-TCAACTTATATACTGATGGAAGTTCCT 179

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|||
Db        180 TTGTAGAAAAAGGACTTCGAAAAGCGGGGTATGCAGTGGTCAGTGATAATGGAATACTTG 239

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              |||||||||||||||||||||||||| |||:||||:||||||||||||:|||:|:|:
Db        240 ---GTAATCCCCTCACTCCAGGAACTAGCGCTCAGCTAGCAGAACTAATAGCCCTCACTG 296

Qy       4047 GGG-CACTAGAATTAGGAGAA-GRAAAAAGGGYAAATATATATACAGACTCTRARTATGC 4104
              ||| |||||||||||||||| |:|||| |||:|||| ||||||||||||:|:|||||
Db        297 GGGCCACTAGAATTAGGAGAAGGGAAAATGGGTAAATACATATACAGACTCTAAGTATGC 356

Qy       4105 TYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGA 4164
              |:|||||||| ||||||||||::|||||||:|:|||||||||||| ||||||||:||
Db        357 TTACCTAGTCCTCCATGCCCATGCAGCAATATGGAGAGAAAGGGAATTGCTAACTTCCGA 416

Qy       4165 GRGAACACCTAT 4176
              |:||||||||||
Db        417 GGGAACACCTAT 428


RESULT 880
CQ741551
LOCUS       CQ741551                409 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 27485 from Patent WO02068579.
ACCESSION   CQ741551
VERSION     CQ741551.1  GI:42349833
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
```

```
  JOURNAL   Patent: WO 02068579-A 27485 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             4.6%;  Score 341.4;  DB 1;  Length 409;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  376;  Conservative    6;  Mismatches   27;  Indels    4;  Gaps    2;

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
             |||||||||||||||||||||||||||||| |||| |||:||||||||||||||||||||
Db          1 AATCTATGTGCTTCCTCTCATTCTTAGTGTCCCC-ATGACCATCTACACTGAACAAGATT 59

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
             |||||||| |   | || ||||||||||||||||||||||||||||||||||||||||||
Db         60 TATACAGTCA---CGTACCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 116

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTC 6613
             | |||||||||||||||| | |||||||||||||||||||| |||||||||||||| |||
Db        117 TTGGAGCAGGAGTGCTAGGCGGACTAGGTACTGGCATTGGCAGTATCACAACCTCTGCTC 176

Qy       6614 AGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCC 6673
             |||||||||||||||||||||||||||||||||  |||||||| ||||  | |||||||
Db        177 AGTTCTACTACAAACTATCTCAAGAACTAAATGGTAACATGGAATGGGTTACTGACTCCC 236

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
             |||||||||||||| |||||||||||||||||||||||||||||:||||| |||||
Db        237 TGGTCACCTTGCAAGTTCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCAAAGAG 296

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
             |||||||||:||||||| ||||:|| ||||||| |||||||||| |||||| |||||||||
Db        297 CTTTAGACTTGCTAACCACTGAAAGTGGGGGAATCTGTTTATTATTAGGGCAAGAATGCT 356

Qy       6794 GTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAAATTCSA 6846
             |||||||||| |||||||| |||||| | ||||||||||||:||||||| :|
Db        357 GTTATTATGTAAATCAATCTGGAATCATGACTGAGAAAGTTAAAGAAATTTGA 409


RESULT 881
AR344410
LOCUS       AR344410                 429 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 136 from patent US 6582703.
ACCESSION   AR344410
VERSION     AR344410.1  GI:33740351
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 136 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             4.6%;  Score 341;  DB 1;  Length 429;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches  360;  Conservative   39;  Mismatches   29;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
             ||||||||:||:|::|||||||||||||| |||||| ||||:|||:| |||||| ||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATGGGGATGACTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||| |||||||||||||||:|||||:|||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGCGCTCTTAAATTTCCTT 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          |||||||  |:|:|||||  ||||||||||  |||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAAGGCTCACCTCTGCTCACACCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| |||||:|||||||| |||||:::|||||||  |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCACCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||||||||||||||:|| ||||||||||||||||||:|::|||||||||:|| :|
Db        232 ATACTGGCTTATCCTCATCCCATAACCCTAAAGCAACTAAGAGGGTTCCTTGGCATATCA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:||||||| ||| | |:  | :|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATT-CCCGGATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||| ||:|||||||||:|:||||| :::||||
Db        351 AATTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAAACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 882
AB055878
LOCUS       AB055878                 374 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:U937-1.
ACCESSION   AB055878
VERSION     AB055878.1  GI:15207787
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 374)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="U937-1"
                     /db_xref="taxon:9606"
                     /cell_line="U-937"
     LTR             1. .374
                     /note="HERV-W family"

  Query Match             4.5%;  Score 340.6;  DB 1;  Length 374;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  361;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120
```

```
Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACC- 386
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCA 240

Qy        387 CACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 446
              |     |||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGCTCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 300

Qy        447 CCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGG 506
              ||||||||:||| |||||||||||||||||||||:|||||||||||||||||||||||
Db        301 CCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGG 359

Qy        507 CCTGCTACCRTCTTG 521
              |||||||||:|||||
Db        360 CCTGCTACCATCTTG 374


RESULT 883
AB066658
LOCUS       AB066658                372 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HFB-10.
ACCESSION   AB066658
VERSION     AB066658.1  GI:18181900
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HFB-10"
                     /db_xref="taxon:9606"
     LTR             1. .372
                     /note="HERV-W LTR family"

  Query Match             4.5%;  Score 339.6;  DB 1;  Length 372;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches  360;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||   |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          1 CGAAAAGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGT 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||||||:|||||||||||| ||||||||||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCAT-TTGGAAGCGGC 358

Qy        508 CTGCTACCRTCTTG 521
              |||||||||:|||||
Db        359 CTGCTACCATCTTG 372


RESULT 884
AB079275
LOCUS       AB079275                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-7.
ACCESSION   AB079275
VERSION     AB079275.1  GI:28201106
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-7"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             4.5%;  Score 339.5;  DB 1;  Length 373;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   14;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 NGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCCG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||||| |||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 GGTTTGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:|||||||||||| |||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||| ||||:|||||||||||||||||||| ||
```

```
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGTGGT 359

Qy        508 CTGCTACCRTCTT 520
                | |||:||||
Db        360 TCACCACCATCTT 372


RESULT 885
AX407891/c
LOCUS       AX407891               410 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 538 from Patent WO0229103.
ACCESSION   AX407891
VERSION     AX407891.1  GI:21440596
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 538 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .410
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. AA250958"

  Query Match             4.5%;  Score 339;  DB 1;  Length 410;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches  372;  Conservative    8;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       7173 CCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAG 7232
              ||||| |||||||||||||||||||| |||: || ||||||||| |||||||||||| |
Db        410 CCCCAGTTCAGCAGGAAGCAGTTAGAGTGGTCATCAGCCAACCTCACCAACAGCACTTGG 351

Qy       7233 GTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTA 7292
              |||||||||||||||  |||||||||||||||||||||||||||||||||||||||| ||:||
Db        350 GTTTTCCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGATTA 291

Qy       7293 AGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAA 7352
              |||||||:||| |||||:|||||||||||||| |||||||||||||||||||||||||||||
Db        290 AGAATCCCTAAACCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAA 231

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
              |||||:|||||||  |||||||||||||||||||||||||||||||||||||||||| ||||||
Db        230 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 171

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGA-GAGCACAGCAGGAGGGACAATGATCGG 7471
              |||||||||||||||||||||:||||||:||| ||||||||  |||||||||| ||| |
Db        170 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGGAGCACAGTGGGAGGGACAAGGATTGC 111

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGC-AACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||  |  |||||||| || |||||||||| |||||||||||| ||||||||
Db        110 AATATAAACCCAGGCATTCGAGCCAGCGAACGGCAACCGCCTTTGGGTCCCTTCCCTTTG 51

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTG 7580
              |||||||||||||||||||  |||||||||||||||||||||||||:|||
Db         50 TATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTG 1


RESULT 886
AR344411
LOCUS       AR344411               429 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 137 from patent US 6582703.
ACCESSION   AR344411
VERSION     AR344411.1  GI:33740352
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 137 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             4.5%;  Score 339;  DB 1;  Length 429;
  Best Local Similarity   81.7%;  Pred. No. 0;
  Matches  358;  Conservative   40;  Mismatches   30;  Indels   10;  Gaps    2;

Qy       2949 GACTTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              ||||||||:||:|::|||||||||||||| |||||| ||||:|||:  |||||||||| ||
Db          1 GACTTGAGCCAGTCCTCATACCTGGACATTCTTGTTCTTCAGTATAGGGATGATTTAATT 60

Qy       3009 TTRGCYGCCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAAGCRCTCTTMAATTTCCTC 3068
              |:||: ||::|||||||||||||||| |||||||||||||||| :|||||:|||||||||
Db         61 ATAGCCACCCATTCAGAAACCTTGTGGCATCAAGCCACCCAAGTGCTCTTAAATTTCCTC 120

Qy       3069 GCYACCTGTGGCTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGCAGGTTAAA 3128
              ||:|||||||||          ||||||| |: :|||||:|||||||||||||||||||||
Db        121 GCTACCTGTGGC---------TCCAAACAAAGGGCTCAGCTCTGCTCACAGCAGGTTAAA 171

Qy       3129 TACTTAGGRCTAARATTATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYATCCAGCCT 3188
              ||||||||:||||:|||||||||||| | |||:|||||||| ||||||:::|||||||| |
Db        172 TACTTAGGGCTAAAATTATCCAAAGTCGCCAGGGCCCTCAGAGAGGAACGTATCCAGCGT 231

Qy       3189 ATACTGGCTTATCCTCATCYCAAAACCCTAAAGCAACTAAGRGRRTTCCTTGGCRTAAYA 3248
              ||||||| ||||||||||:|||||||  |||||||||||||:|::|||||||||:|||:|
Db        232 ATACTGGATTATCCTCATCCCAAAACCATAAAGCAACTAAGAGGGTTCCTTGGCATAACA 291

Qy       3249 GGYTTCTGCCGAAWATGGATTCCCCAGGTWTGGCRAAATAGCCAGGYCATTAWATACAST 3308
              | :||||||||||:|||||||||||| | |:  | :|||||||||||:|||||: |||| |
Db        292 GCCTTCTGCCGAATATGGATTCCCC-GATACAGTGAAATAGCCAGGCCATTATGTACATT 350

Qy       3309 AATTAAGGAAACTCAGAAAGCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAAGT 3368
              |||||||||||||||||||||||||||||||| ||:|||||||:|:|||| :::||||
Db        351 AGTTAAGGAAACTCAGAAAGCCAATACCCATATAGTAAGATGGACACCTGAGACAGAAGT 410

Qy       3369 GGCTTTCCAGGCCCCTAA 3386
              |||||||||||||  ||
Db        411 GGCTTTCCAGGCCCTAAA 428


RESULT 887
AB064532
LOCUS       AB064532               373 bp    DNA     linear   PRI 05-JAN-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR in Xq21.3 region, isolate:X7-1.
ACCESSION   AB064532
VERSION     AB064532.1  GI:18148364
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Isolation of HERV-W LTR in Xq21.3 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82-51-510-3346,
```

```
          Fax:82-51-581-2962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="X7-1"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq21.3"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.5%;  Score 338.3;  DB 1;  Length 373;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   15;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCATCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||| |||||||||| ||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCTTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| |||||||||| ||| ||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||||| ||||:||||||||||||||||||||||
Db        301 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              |||| | |:|||||
Db        360 CTGCCAACATCTTG 373


RESULT 888
AB071973
LOCUS       AB071973                373 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch4.
ACCESSION   AB071973
VERSION     AB071973.1  GI:19698208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch4"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             4.5%;  Score 338.3;  DB 1;  Length 373;
  Best Local Similarity   94.4%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   15;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCGGCACGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||||||||||||||||||| |||||||||||| |||||||||  ||||||
Db        181 GGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCTTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| |||||||||| ||| ||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||||| ||||:||||||||||||||||||||||
Db        301 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              |||| | |:|||||
Db        360 CTGCCAACATCTTG 373


RESULT 889
AB079279
LOCUS       AB079279                389 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-4.
ACCESSION   AB079279
VERSION     AB079279.1  GI:28201110
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-4"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>389
                     /note="HERV-W LTR like element"

  Query Match             4.5%;  Score 337.9;  DB 1;  Length 389;
  Best Local Similarity   93.3%;  Pred. No. 0;
  Matches  364;  Conservative    5;  Mismatches    4;  Indels   17;  Gaps    2;
```

```
Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        241 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||||||||:||| ||||||||||||||||||:||||||||
Db        301 TCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||:|||||
Db        360 CCATCTTGGAAGCGGCCTGCTACCATCTTG 389


RESULT 890
AB079269
LOCUS       AB079269                389 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-1.
ACCESSION   AB079269
VERSION     AB079269.1  GI:28201100
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-1"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>389
                     /note="HERV-W LTR like element"

  Query Match             4.5%;  Score 336.3;  DB 1;  Length 389;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  363;  Conservative    5;  Mismatches    5;  Indels   17;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        241 TTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGAT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||||||||||:||| ||||||||||||||||||||:||||||||
Db        301 TCCATTCCTTGGAATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||:||| |
Db        360 CCATCTTGGAAGCGGCCTGCTACCATCTGG 389


RESULT 891
AB021920
LOCUS       AB021920                434 bp    DNA     linear   VRL 26-JAN-2000
DEFINITION  Human endogenous retrovirus W gene, clone HWX3, similar to MSRV
            polyprotein sequence.
ACCESSION   AB021920
VERSION     AB021920.1  GI:4996594
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (sites)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Identification and phylogeny of novel human endogenous retroviral
            sequences belonging to the HERV-W family on the human X chromosome
  JOURNAL   Arch. Virol. 144 (12), 2403-2413 (1999)
   PUBMED   10664393
REFERENCE   2  (bases 1 to 434)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1998) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:87786"
                     /clone="HWX3"
                     /note="HERV-W on human X chromosome"
     misc_feature    1. .434
                     /note="similar to MSRV polyprotein sequence in Acc#
                     AF009668"

  Query Match             4.5%;  Score 336;  DB 1;  Length 434;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  372;  Conservative   26;  Mismatches   28;  Indels   12;  Gaps    3;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||| |||||||||||||||||||||:| ||||||||||||||||||||
Db          1 AGTGCTGCAACTGCACACTTGTGCAACTCTTAACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:|||||||||||||| :| ||| || |||| || ||| || |||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA------ 3920
              ||||:||||:|||||:||||||||||:|||| |||||||||| ||||||||||
Db        121 CCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTT 179
```

```
Qy       3921 --GTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGG 3978
              |||||||  ||||||||||||||||  |||||:|||||  ||||||||||||||||||||
Db        180 TGGTTCCTTGGTAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGG 239

Qy       3979 AATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGC 4038
              ||||:|||||||||||||||  |||||||  ||||||  |||:||||:||||||  |||||||
Db        240 AATACTTGAAAGTAATCCCCTTACTCCAGAAACTAGCGCTCAGCTGGCAGAATTAATAGC 299

Qy       4039 CYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRA 4098
              |:|||:| ||||||||  || |    |||||:||||||||:|||| ||||||||||||||:|
Db        300 CCTCACTCGGGCACTAAAACT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAA 356

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
              :||||||:| ||||||  ||||||||  |  ::||||||||:|:||||||||||||||||||
Db        357 GTATGCTTATCTAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTAAC 416

Qy       4159 TTCYGAGRGAACACCTAT 4176
              |||:|||:||||||||||
Db        417 TTCCGAGGGAACACCTAT 434


RESULT 892
CQ052505/c
LOCUS       CQ052505                405 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 3325 from Patent WO0157270.
ACCESSION   CQ052505
VERSION     CQ052505.1  GI:41027011
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 3325 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN HBL100, SIGNAL = 1.2"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| ||||||||||:||||||||||:||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||:|||| ||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 893
CQ067587/c
LOCUS       CQ067587                405 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 3387 from Patent WO0157278.
ACCESSION   CQ067587
VERSION     CQ067587.1  GI:41037456
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 3387 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN HELA, SIGNAL = 1.4"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| ||||||||||:||||||||||:||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||:|||| ||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 894
CQ094637/c
LOCUS       CQ094637                405 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3496 from Patent WO0157272.
ACCESSION   CQ094637
```

```
VERSION     CQ094637.1  GI:41063663
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 3496 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN PLACENTA, SIGNAL = 3"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::||||||||  |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:|||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||| |||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              ||||||||||||||| ||||||||||:||||||||||:|||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:||||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 895
CQ133399/c
LOCUS       CQ133399                405 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3421 from Patent WO0157276.
ACCESSION   CQ133399
VERSION     CQ133399.1  GI:41090755
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 3421 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.3"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::||||||||  |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:|||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||| |||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              ||||||||||||||| ||||||||||:||||||||||:|||||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:||||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 896
CQ171968/c
LOCUS       CQ171968                405 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3364 from Patent WO0157274.
ACCESSION   CQ171968
VERSION     CQ171968.1  GI:41166704
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 3364 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN HEART, SIGNAL = 1"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342
```

```
Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:|||||||||  ||||||||||::||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||||:|||| ||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 897
CQ201114/c
LOCUS       CQ201114                405 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3391 from Patent WO0157271.
ACCESSION   CQ201114
VERSION     CQ201114.1  GI:41186778
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 3391 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN BT474, SIGNAL = 1.2"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:|||||||||  ||||||||||::||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
             |:||||||||||||:|||| ||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 898
CQ216631/c
LOCUS       CQ216631                405 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3470 from Patent WO0157273.
ACCESSION   CQ216631
VERSION     CQ216631.1  GI:41197955
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 3470 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.5"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
             ||||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
             |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
             ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
             :|:||||||||||:|||||||||  ||||||||||::||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
             |||||||    |||||||||||| ||||||||||||||||:|||||::|||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
             |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47
```

```
Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||||:|||| |||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 899
CQ255195/c
LOCUS       CQ255195                405 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3456 from Patent WO0157277.
ACCESSION   CQ255195
VERSION     CQ255195.1  GI:41227673
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 3456 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.4"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db         400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:|||||||:|||||||||| ||||:|||:|||
Db         341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db         282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db         222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||| ||||||||||||||:||||::|||||||||||||||
Db         162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db         106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||||:|||| |||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 900
CQ292282/c
LOCUS       CQ292282                405 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3387 from Patent WO0186003.
ACCESSION   CQ292282
VERSION     CQ292282.1  GI:41252859
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 3387 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC010778.2
                     EXPRESSED IN LUNG, SIGNAL = 1.5"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:||||||||||||:||||||||:||||||||||||
Db         400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:|||||||:|||||||||| ||||:|||:|||
Db         341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||||::||||::|||||| ||| |||| ||||||||| ||||||||||
Db         282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db         222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||| ||||||||||||||:||||::|||||||||||||||
Db         162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| ||||||||||:||||||||||||:||||||||||||||||||
Db         106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||||:|||| |||||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 901
CQ329310/c
LOCUS       CQ329310                405 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3404 from Patent WO0157275.
ACCESSION   CQ329310
VERSION     CQ329310.1  GI:41277877
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 3404 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
                    /note="MAP TO AC010778.2
                    EXPRESSED IN BRAIN, SIGNAL = 1.1"

  Query Match             4.5%;  Score 335.8;  DB 1;  Length 405;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches  358;  Conservative   30;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       3381 CCCTAAAGAAGGCCTTAAACCCAAGYCCCAGTGTTAAGYTTGCCAACRGGGCAAGACTTT 3440
              |||||||||||||| | ||||||||:|||||||||||||:|||||||||:||||||||||||
Db        400 CCCTAAAGAAGGCCCT-AACCCAAGCCCCAGTGTTAAGCTTGCCAACAGGGCAAGACTTT 342

Qy       3441 TSTTYATAYRTCACAGAAAAAAACAGRAAYAGCTCTRGGAGTCCTTACACAGRTCCRAGG 3500
              |:||:|||::|||||||| |||| ||:||:||||||:|||||||||| ||||:|||:|||
Db        341 TCTTTATATGTCACAGAAGAAAA-AGGAATAGCTCTAGGAGTCCTTATACAGGTCCGAGG 283

Qy       3501 GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTGATGTAGTGGCAAAGGGTTG 3560
              ||:||||||||||::||||::|||||| ||| |||| |||||||| ||||||||||
Db        282 GACGAGCTTGCAACCTGTGGCATACCTGAATAAAGAAAATGATGTAGTAGCAAAGGGTTG 223

Qy       3561 RCYTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTATCTGAAGCAGTTAAAAT 3620
              :|:||||||||||:|||||||||  ||||||||||::|||||||||||||||||||||||
Db        222 GCCTCATTGTTTACGGGTAGTGGCAGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAT 163

Qy       3621 AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGC 3680
              |||||||    |||||||||| |||||||||||||||:|||||::||||||||||||||
Db        162 AATACAG----GAGATCTTACTCTGTGGACATCTCATGATGTGAACAGCATACTCACTGC 107

Qy       3681 TAAAGGAGACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGA 3740
              |||||||||||||||| |||||||||||:||||||||||||:||||||||||||||||||
Db        106 TAAAGGAGACTTGTGGATGTCAGACAACCGTTTACTTAAATATCAGGCTCTATTACTTGA 47

Qy       3741 ARGGCCAGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYC 3786
              |:|||||||||||:|||| |||||||||||||||||||||||:|
Db         46 AGGGCCAGTGCTGCAACTGCGCACTTGTGCAACTCTTAACCCAGCC 1


RESULT 902
AB079273
LOCUS       AB079273                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-5.
ACCESSION   AB079273
VERSION     AB079273.1  GI:28201104
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-5"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             4.5%;  Score 335.7;  DB 1;  Length 373;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches  351;  Conservative    5;  Mismatches   16;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db         61 CGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||| |
Db        121 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCCCG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||||| |||||||||||||||||| ||||||| |||||||||||||||||||||||
Db        181 GGTTTGTCCTAATTGAGCTGAACACTAGTCACTGGGAGCCATGGTTCTCTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:||||||||||| ||||||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAACTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||||||||||||||||| ||||:|||||||||||||||||||| ||
Db        301 CGTGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGTGGT 359

Qy        508 CTGCTACCRTCTT 520
                 | |||:||||
Db        360 TCACCACCATCTT 372


RESULT 903
BV193474/c
LOCUS       BV193474                401 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm177987 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV193474
VERSION     BV193474.1  GI:48036401
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 401.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>401

  Query Match             4.5%;  Score 335.1;  DB 1;  Length 401;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  363;  Conservative    4;  Mismatches   33;  Indels    1;  Gaps    1;

Qy       5134 AGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCA 5193
              || |||||| | |||||||||||| |||||||| |||||||||||||||| |||||| |
```

```
Db        401 AGTCCATACAAGATGCTGTATGGATGGCCCTTCCTAACCAATGACCTTGTGTTTGACCAA 342

Qy       5194 -AGACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAA 5252
              ||| |||||||||| ||| |||||||||||||||||| |||| ||||||||||||||||
Db        341 GAGAGAGCCAACTTAATTGAAGACATCACCTCCTTAGCTAAATGTCAACAAGTTCTTAAA 282

Qy       5253 ACATTACAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATG 5312
              |||||||| ||||||| ||||||||| |||||||| ||| ||||||||||: ||||||||
Db        281 ACATTACAGGGAACCTGTCCCTGAGAGGAGGGAAAGGAATTATTCCACCCTGGTGACATG 222

Qy       5313 GTATTAGTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCC 5372
              ||||||||||||||||||::|||| ||||||||||||||||||||||||||||||||||
Db        221 GTATTAGTCAAGTCCCTTCCYTCTAGTTCCCCATCCCTAGATACATCCTGGGAAGGACCC 162

Qy       5373 TACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATA 5432
              |||||||||||||||:|||||||||   ||||||||||||||||||||||||||||||||
Db        161 TACCCAGTCATTTTATCTACCCCAACCATGGTTAAAGTGGCTGGAGTGGAGTCTTGGATA 102

Qy       5433 CATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAAC 5492
              |||||| ||  ||||||| |||||||||||||||||| || ||||||||||||||||
Db        101 CATCACGCTCAAGTCAAACCCTGGATACTGCCAAAGGAGCCCGAAAATCCAGGAGACAAT 42

Qy       5493 GCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCT 5533
              ||| | |||||||||||||||||||||||| ||| ||||||
Db         41 GCTGGGTATTCCTGTGAACCTCTAGAGGATCTGCACCTGCT 1
```

RESULT 904
AB024477
LOCUS AB024477 426 bp DNA linear PRI 21-OCT-1999
DEFINITION Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV, isolate:GIW4-5.
ACCESSION AB024477
VERSION AB024477.1 GI:6063508
KEYWORDS .
SOURCE Hylobates agilis (agile gibbon)
ORGANISM Hylobates agilis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (sites)
AUTHORS Kim,H.S., Takenaka,O. and Crow,T.J.
TITLE Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates
JOURNAL J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
PUBMED 10573154
REFERENCE 2 (bases 1 to 426)
AUTHORS Kim,H.-S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital, Psychiatry; Headington, Oxford OX3 7JX, U.K.
(E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397, Fax:441865244990)
FEATURES Location/Qualifiers
source 1. .426
/organism="Hylobates agilis"
/mol_type="genomic DNA"
/isolate="GIW4-5"
/db_xref="taxon:9579"
/sex="Male"
/tissue_type="Blood"
misc_feature <1. .>426
/note="similar to pol gene of human endogenous retrovirus family (HERV-W) and multiple sclerosis-associated retrovirus (MSRV)"

Query Match 4.5%; Score 334.1; DB 1; Length 426;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 366; Conservative 24; Mismatches 25; Indels 5; Gaps 3;

```
Qy       3757 CTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAAGAAAAGATA 3816
              |||  |||||||||||||||||||||||:|  || |||||||||||||||||||||||||||
Db          12 CTGCTCACTTGTGCAACTCTTAACCCAGCCACACTTCTTCCAGACAATGAAGAAAAGATA 71

Qy       3817 RAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGACCTTYTAGAR 3876
              :||:|||||||||||||| :  ||| |||||||||| ||||||| ||||||||||:||||:
Db          72 GAACATAACTGTCAACAGGGGATTGCTCAAACCTACGCCACTCAAGGGGACCTTCTAGAG 131

Qy       3877 GTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAA 3936
              |||||:||||||||||:|||| ||||||||||||||||||||| ||||||||||||||||
Db         132 GTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAATTTCCTTTGTAGAAAA 190

Qy       3937 AGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTGAAAGTAATCC 3996
              |||||||||||||||:|||||||||||| |||| ||||||||||||||:|||||||||||
Db         191 AGGACTTCGAAAAGCGGGGTATGCAGCGGTCCGTGATAATGGAATACTTGAAAGTAATCC 250

Qy       3997 CCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTKGGGCACTAGA 4056
              |||||||||||||||||  ||:||||:|||||||||||||:|||: :|||||| | |
Db         251 CCTCACTCCAGGAACTAGCACTCAGCTGGCAGAACTAATAGCCTTCACCGGGGCACCAAA 310

Qy       4057 ATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTYACCTAGTCNT 4116
              |||   |||||:|||||||| |||||  |||||||||||| |:||||||:|||||||| |
Db         311 ATT---AGAAGGAAAAAGGGGAAATAAGTATACAGACTCT-AATATGCTTACCTAGTCCT 366

Qy       4117 CCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGRGAACACCTAT 4176
              |||||||||  ::|||||||:|: |||||||||||||||||||:|||:|||||||||||
Db         367 CCATGCCCACACAGCAATATGGAGTGAAAGGGAATTCCTAACTTCTGAGGGAACACCTAT 426
```

RESULT 905
AB024478
LOCUS AB024478 415 bp DNA linear PRI 21-OCT-1999
DEFINITION Hylobates agilis DNA, similar to pol gene of HERV-W and MSRV, isolate:GIW4-6.
ACCESSION AB024478
VERSION AB024478.1 GI:6063511
KEYWORDS .
SOURCE Hylobates agilis (agile gibbon)
ORGANISM Hylobates agilis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (sites)
AUTHORS Kim,H.S., Takenaka,O. and Crow,T.J.
TITLE Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates
JOURNAL J. Gen. Virol. 80 (Pt 10), 2613-2619 (1999)
PUBMED 10573154
REFERENCE 2 (bases 1 to 415)
AUTHORS Kim,H.-S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Heui-Soo Kim, Warneford Hospital, Psychiatry; Headington, Oxford OX3 7JX, U.K.
(E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397, Fax:441865244990)
FEATURES Location/Qualifiers
source 1. .415
/organism="Hylobates agilis"
/mol_type="genomic DNA"
/isolate="GIW4-6"
/db_xref="taxon:9579"
/sex="Male"
/tissue_type="Blood"
misc_feature <1. .>415
/note="similar to pol gene of human endogenous retrovirus family (HERV-W) and multiple sclerosis-associated retrovirus (MSRV)"

Query Match 4.5%; Score 333.5; DB 1; Length 415;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 368; Conservative 25; Mismatches 22; Indels 15; Gaps 3;

```
Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|| |||||||||||||||||||     ||:| |||||||||| |||||||
Db           1 AGTGCTGCGACCGTGCACTTGTGCAACTCTT-----AGCCACATTTCTTCCACACAATGA 55
```

```
Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:| : |||| ||||||||:||||| |||||||||| ||| ||||||||||
Db         56 AGAAAAGATAGACCGTAACCGTCAACAAGTAATTGCTCAAACCTACGCCGCTCGAGGGGA 115

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCT 3926
              ||||:||||:|||||:||||||||||: ||| |||||||||||||||||||||||||||
Db        116 CCTTTTAGAGGTTCCCTTGACTGATCCCAACC-TCAACTTGTATACTGATGGAAGTTCCT 174

Qy       3927 TTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGAATAYTTG 3986
              |||||||||||||||||| |||||:|||||||||||| ||||||||||||||||||:|||
Db        175 TTGTAGAAAAAGGACTTCAAAAAGCGGGGTATGCAGTAGTCAGTGATAATGGAATACTTG 234

Qy       3987 AAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCCYTCAYTK 4046
              ||||||||||||||||||| ||||||||||:||||:||| |||          |||:|
Db        235 AAAGTAATCCCCTCACTCCAAGAACTAGTGCTCAGCTGGCAAAAC---------TCACTC 285

Qy       4047 GGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRARTATGCTY 4106
              ||||||||||||||||||||:|||| |||:||||| ||||||||||||:|:||||||:
Db        286 GGGCACTAGAATTAGGAGAAGGAAAACGGGTAAATAGATATACAGACTCTAAGTATGCTT 345

Qy       4107 ACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAACTTCYGAGR 4166
              |||||||| ||||||||||::| ||||||:|:||||||||||||||||||||:|| :
Db        346 ACCTAGTCCTCCATGCCCATGCAGTAATATGGAGAGAAAGGGAATTCCTAACTTCCGAAG 405

Qy       4167 GAACACCTAT 4176
              ||| | ||||
Db        406 GAATATCTAT 415


RESULT 906
AB021919
LOCUS       AB021919                434 bp    DNA     linear   VRL 26-JAN-2000
DEFINITION  Human endogenous retrovirus W gene, clone HWX1, similar to MSRV
            polyprotein sequence.
ACCESSION   AB021919
VERSION     AB021919.1  GI:4996593
KEYWORDS    .
SOURCE      Human endogenous retrovirus W
  ORGANISM  Human endogenous retrovirus W
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (sites)
  AUTHORS   Kim,H. and Crow,T.J.
  TITLE     Identification and phylogeny of novel human endogenous retroviral
            sequences belonging to the HERV-W family on the human X chromosome
  JOURNAL   Arch. Virol. 144 (12), 2403-2413 (1999)
   PUBMED   10664393
REFERENCE   2  (bases 1 to 434)
  AUTHORS   Kim,H.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1998) Heui-Soo Kim, Warneford Hospital,
            Psychiatry; Headington, Oxford OX3 7JX, U.K.
            (E-mail:heui-soo.kim@psychiatry.oxford.ac.uk, Tel:441865226397,
            Fax:441865244990)
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Human endogenous retrovirus W"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:87786"
                     /clone="HWX1"
                     /note="HERV-W on human X chromosome"
     misc_feature    1. .434
                     /note="similar to MSRV polyprotein sequence in Acc#
                     AF009668"

  Query Match             4.4%;  Score 326.4;  DB 1;  Length 434;
  Best Local Similarity   83.6%;  Pred. No. 0;
  Matches  366;  Conservative   26;  Mismatches   34;  Indels   12;  Gaps    3;

Qy       3747 AGTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGA 3806
              ||||||||:|||| ||||||||||||||| ||||||:| |||||||||||||||||||
Db          1 AGTGCTGCAACTGCACACTTGTGCAACTCTTGACCCAGCCACATTTCTTCCAGACAATGA 60

Qy       3807 AGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGA 3866
              ||||||||||:||:||||||||||||| :| ||| || |||| || ||| || |||||||
Db         61 AGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTTAAACATACGCCGCTTGAGGGGA 120

Qy       3867 CCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA------ 3920
              ||||:| ||:|||||:||||||||||:|||| |||||||||| ||||||||||
Db        121 CCTTCTGGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTACACTGATGGAAGCTCTT 179

Qy       3921 --GTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGG 3978
                |||||| |||||||||||||||| |||||:||||| ||||||||||||||||||||||
Db        180 TGGTTCCTTGGTAGAAAAAGGACTTCAAAAAGTGGGGTGTGCAGTGGTCAGTGATAATGG 239

Qy       3979 AATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGC 4038
              ||||:|||||||| |||||| ||||| | |||||| |||:||||:|||| | ||| |||
Db        240 AATACTTGAAAGTAGTCCCCTTACTCCGGAAACTAGCGCTCAGCTGGCAGGATTAACAGC 299

Qy       4039 CYTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGGGYAAATATATATACAGACTCTRA 4098
              |:|||:| |||||||| ||||   |||||:|||||||:|||| |||||||||||||:|
Db        300 CCTCACTCGGGCACTAAAATT---AGAAGAAAAAAGGGTAAATGTATATACAGACTCTAA 356

Qy       4099 RTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSARAGAAAGGGAATTCCTAAC 4158
              :||||||:| |||||| |||||||| |  ::|||||||:|:|||||||||||||| ||
Db        357 GTATGCTTATCTAGTCCTCCATGCCTACACAGCAATATGGAGAGAAAGGGAATTCCTGAC 416

Qy       4159 TTCYGAGRGAACACCTAT 4176
              |||:|||:||||||||||
Db        417 TTCCGAGGGAACACCTAT 434


RESULT 907
CQ053304/c
LOCUS       CQ053304                409 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 4124 from Patent WO0157270.
ACCESSION   CQ053304
VERSION     CQ053304.1  GI:41027810
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 4124 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN HBL100, SIGNAL = 1.9"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170
```

```
Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 908
CQ068432/c
LOCUS       CQ068432                409 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 4232 from Patent WO0157278.
ACCESSION   CQ068432
VERSION     CQ068432.1  GI:41038301
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 4232 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN HELA, SIGNAL = 2.2"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||||| ||||| ||||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 909
CQ095502/c
LOCUS       CQ095502                409 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4361 from Patent WO0157272.
ACCESSION   CQ095502
VERSION     CQ095502.1  GI:41064528
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 4361 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN PLACENTA, SIGNAL = 3.5"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||||| ||||| ||||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 910
CQ134252/c
LOCUS       CQ134252                409 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4274 from Patent WO0157276.
ACCESSION   CQ134252
VERSION     CQ134252.1  GI:41091608
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 4274 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.8"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
               |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              ||||||||  |||||||| |||||| ||||||||||||||||  | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              |||||  |||||||  ||||||||||||||  |  ||||||||  ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              ||||||||  ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 911
CQ172757/c
LOCUS       CQ172757                409 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4153 from Patent WO0157274.
ACCESSION   CQ172757
VERSION     CQ172757.1  GI:41167493
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 4153 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN HEART, SIGNAL = 1.7"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
               |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230


RESULT 912
CQ201948/c
LOCUS       CQ201948                409 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4225 from Patent WO0157271.
ACCESSION   CQ201948
VERSION     CQ201948.1  GI:41187612
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 4225 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN BT474, SIGNAL = 1.9"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| |||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
               |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230
```

```
Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||   | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| ||||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| ||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 913
CQ217505/c
LOCUS       CQ217505                409 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4344 from Patent WO0157273.
ACCESSION   CQ217505
VERSION     CQ217505.1  GI:41198829
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 4344 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.1"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||| ||||| ||||||||||||||||||||  ||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || ||||||||||||||||||||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||   | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| ||||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| ||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 914
CQ256070/c
LOCUS       CQ256070                409 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4331 from Patent WO0157277.
ACCESSION   CQ256070
VERSION     CQ256070.1  GI:41228550
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 4331 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.9"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              |||||||||||||||||||||| ||||| ||||||||||||||||||||  ||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || ||||||||||||||||||||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||   | ||||| ||||| ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              ||||| ||||||| ||||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| ||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||   | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 915
CQ293157/c
LOCUS       CQ293157                409 bp    DNA     linear   PAT 23-JAN-2004
```

```
DEFINITION  Sequence 4262 from Patent WO0186003.
ACCESSION   CQ293157
VERSION     CQ293157.1  GI:41253734
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 4262 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN LUNG, SIGNAL = 3.6"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| ||||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | |||||  |||||  ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              |||||  |||||||  |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||    | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 916
CQ330127/c
LOCUS       CQ330127                409 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4221 from Patent WO0157275.
ACCESSION   CQ330127
VERSION     CQ330127.1  GI:41278694
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 4221 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AF121782.1
                     EXPRESSED IN BRAIN, SIGNAL = 1.9"

  Query Match             4.3%;  Score 324.7;  DB 1;  Length 409;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  358;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        110 TTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCG 169
              ||||||||||||||||||||||| ||||| ||||||||||||||||||||||  |||||||
Db        409 TTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCACCG 350

Qy        170 TCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCC 229
              |||||||||||||||||||||| |||| ||| |||||||||| |||| ||||  |||||
Db        349 TCCACCACTGCTGTTTGCCACCATCGCAGACCCGCCGCTGACTTCCATTCCTCCAGATCC 290

Qy        230 TGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGC 289
              |||| || || |||||||||||||||||| ||:|||||||||||| || |||||  |||
Db        289 GGCAGAGTATCTGCTGTGCTCCTGATCCAGTGAGGCGCCCATTGCCACTTCCAATCAGGC 230

Qy        290 TAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAA 349
              |||||||| |||||||| |||||| |||||||||||||||||  | |||||  |||||  ||
Db        229 TAAAGGCTCGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATCCTAATCGAGCTAAA 170

Qy        350 CACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAA 409
              |||||  |||||||  |||||||||||||| |  |||||||| ||||||||||:| ||||||
Db        169 CACTAGTCACTGGCTTCCATGGTTCTCTCCCATGACCCACAGCTTCTAATAGAGCTATAA 110

Qy        410 CACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGG 469
              |||| ||| |||||||||||||||||||||||||||||||| |||||||:||| |||| |||||
Db        109 CACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATTCGTGAGGCCAA-GAACCCCAGG 51

Qy        470 TCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||||||| ||||:||||||| |||||||||||| |||    | |||:|||
Db         50 TCAGAGAACACGAGGCTTGCCGCCATCTTGGAAGTGGCTCACCACCATCT 1


RESULT 917
BV193437/c
LOCUS       BV193437                401 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm177856 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV193437
VERSION     BV193437.1  GI:48036346
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 401.
FEATURES             Location/Qualifiers
```

```
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>401

  Query Match             4.3%;  Score 324.3;  DB 1;  Length 401;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  351;  Conservative    5;  Mismatches   35;  Indels    1;  Gaps    1;

Qy         95 TCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTG 154
              |||||||||| |||||||||||||||||||||||||| ||||| ||||||||| |||||
Db        401 TCACTCTATTCAATCTTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTCAAGCTG 342

Qy        155 AGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCC 214
              |||||||||||||| |||||||||||||||||||  ||   ||| ||| | |||||||| |
Db        341 AGCTTTTGCTCACCATCCACCACTGCTGTTTGCCGCCGTTGCAGACCCGTCGCTGACTTC 282

Qy        215 CATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGC 274
              ||||||||  ||||  |||||||||||| ||||||||||||||||||||:|| ||||||||
Db        281 CATCCCTCCAGATCTGGCAGGGTGTCCACTGTGCTCCTGATCCAGCGAGGCACCCATTGC 222

Qy        275 CGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGT 334
              |||||||||| |||||||||:|||||||||||| ||||||| ||||||||||||||| ||
Db        221 CGCTCCCAATCGGGCTAAAGSCTTGCCATTGTTCCTGCACAGCTAAGTGCCTGGGTTCGT 162

Qy        335 TCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTT 394
              |||||  |||||||||||| ||||||||||||||||||||||||  ||||||| |||||
Db        161 CCTAATCAAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCCATGACCCATGGCTT 102

Qy        395 CTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGG 454
              ||||||:| |||||||||| ||| ||||||||||||||||||||||||||||||||||||
Db        101 CTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGG 42

Qy        455 SCAACGAACTCCAGGTCAGAGAATACGARGCT 486
              :||| |||| |||||||||||| ||||:|||
Db         41 CCAA-GAACCCCAGGTCAGAGAACACGAGGCT 11


RESULT 918
G34081/c
LOCUS       G34081                   369 bp    DNA     linear   STS 01-JUL-1997
DEFINITION  human STS SHGC-51483, sequence tagged site.
ACCESSION   G34081
VERSION     G34081.1  GI:2227385
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: CTTACTGAGGGCTCTGGTGC
            Primer B: ACCCGTTCAGAAACCTTGTG
            STS size: 159
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                    Template:           25 ng
                    Primer:             each 1 uM
                    dNTPs:              each 200 uM
                    Taq Polymerase:     0.05 units/ul
                    Total Vol:          10 ul

            Buffer:
                    MgCl2:              2.5 mM
                    KCl:                50  mM
                    Tris-HCl:           20  mM
                    pH:                 8.3

            Redeveloped chromosome 4 markers for scoring on  oligonucleotide
            arrays.  Designed and developed at the Stanford Human Genome Center
            on sequences previously developed as STSs at the Stanford Human
            Genome Center or the Whitehead Institute.
FEATURES                Location/Qualifiers
     source             1. .369
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /map="4"
     STS                177. .335
     primer_bind        177. .196
     primer_bind        complement(316. .335)

  Query Match             4.3%;  Score 324.1;  DB 1;  Length 369;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches  322;  Conservative   33;  Mismatches   13;  Indels    1;  Gaps    1;

Qy       2981 TGTCCTTCRGTAKGTGGATGATTTACTTTTRGCYGCCYRTTCAGAAACCTTGTGCCATCA 3040
              | ||||| :|||  |||||||||||| ||||:||: ||::||||||||||||||||||||||
Db        369 TNTCCTTTGGTACATGGATGATTTATTTTTAGCCACCCGTTCAGAAACCTTGTGCCATCA 310

Qy       3041 AGCCACCCAAGCRCTCTTMAATTTCCTCGCYACCTGTGGCTACAWGGTTTCCAAACSARA 3100
              |||||||||||:|||||:||||||||||:|||||||||||||:||||||||||:|:|
Db        309 AGCCACCCAAGCGCTCTTAAATTTCCTCGCCACCTGTGGCTACAAGGTTTCCAAACCAAA 250

Qy       3101 RGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGGRCTAARATTATCCAAAGGCACCAR 3160
              :|||||:|||||||||||||||||||||||||||||:||||:|||||||||||||||||:
Db        249 GGCTCAGCTCTGCTCACAGCAGGTTAAATACTTAGGGCTAAAATTATCCAAAGGCACCAG 190

Qy       3161 GGCCCTCAGTGAGGAAYRYATCCAGCCTATACTGGCTTATCCTCATCYCAAAACCCTAAA 3220
              ||||||||| |||||::: |||||||||||||||||||||||||||:|||||||||||
Db        189 AGCCCTCAGTAAGGAATGTGTCCAGCCTATACTGGCTTATCCTCATCCCAAAACCCTAAA 130

Qy       3221 GCAACTAAGRGRRTTCCTTGGCRTAAYAGGYTTCTGCCGAAWATGGATTCCCCAGGTWTG 3280
              |||||||||:|::|||||||||:|||:|||:|||||||  |:||||||| |||||||:
Db        129 GCAACTAAGAGGGTTCCTTGGCATAACAGGCTTCTGCCATATATGGATT-CCCAGGTACA 71

Qy       3281 GCRAAATAGCCAGGYCATTAWATACASTAATTAAGGAAACTCAGAAAGCCAATACCCATT 3340
              ||:|||||||||:|||||:|||||:|||||||||||||||||||||||||||||||||||
Db          70 GCGAAATAGCCAGGCCATTATATACACTAATTAAGGAAACTCAGAAAGCCAATACCCATT 11

Qy       3341 TARTAAGAT 3349
              ||:||||||
Db          10 TAGTAAGAT 2


RESULT 919
BV193111
LOCUS       BV193111                 401 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm176750 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV193111
VERSION     BV193111.1  GI:48035871
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 401.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>401

  Query Match             4.3%;  Score 321.5;  DB 1;  Length 401;
  Best Local Similarity   87.8%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   43;  Indels    1;  Gaps    1;

Qy         96 CACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGA 155
              |||||||||||||||||||||||||||||||||||||||| ||||| ||| |||| || ||||
Db          1 CACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCGTGTTTGTTATGGCTTGACCTGA 60

Qy        156 GCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCC 215
              ||||| ||||   |||||| ||||||||||||||   | || |||  || |||||| ||
Db         61 GCTTTCGCTCGTTGTCCACGACTGCTGTTTGCCACTGTCACAGACCCACCACTGACTTCC 120

Qy        216 ATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCC 275
              ||| |||||||||| ||||||||| | |||||||||||||||||| |:||||||||||||
Db        121 ATCTCTCTGGATCCGGCAGGGTGTTCACTGTGCTCCTGATCCAGCAAGGCGCCCATTGCC 180

Qy        276 GCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTT 335
              ||||  || ||||||||||:|||||||||| |||||| ||||||||||||||||||  |
Db        181 ACTCCTGATCGGGCTAAAGGYTTGCCATTGTTCCTGCATGGCTAAGTGCCTGGGTTCATC 240

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              ||||| |||||||||||||| |||||||||||| |||||||||| ||||||||||||||||
Db        241 CTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTC 300

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGS 455
              |||||:| |||||||||| |||||||||||||||||||||||||||||||||||||||| |||:
Db        301 TAATAGAGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTCAGGC 360

Qy        456 CAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCT 497
              ||| |||| |||||||||||||| ||||:|||||||| |||||
Db        361 CAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCATCATCT 401


RESULT 920
AB066652
LOCUS       AB066652                389 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HB-15.
ACCESSION   AB066652
VERSION     AB066652.1  GI:18181894
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-15"
                     /db_xref="taxon:9606"
     LTR             1. .389
                     /note="HERV-W LTR family"

  Query Match             4.3%;  Score 320.9;  DB 1;  Length 389;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   15;  Indels   17;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:||||||| ||||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              ||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||| ||||||||||||||| |||||:| |||||||||| ||||||||||||||||
Db        241 TTCTCTTCCGTGACCCACGGCTTNTAATAGAGCTATAACACTCACCACATGGCCCAAGAT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||| |||||:||| |||||||||||||||| ||||:||||||||
Db        301 TCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||:|||||
Db        360 CCATCTTGGAAGCGGCCTGCTACCATCTTG 389


RESULT 921
AB079282
LOCUS       AB079282                389 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-7.
ACCESSION   AB079282
VERSION     AB079282.1  GI:28201113
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
```

```
           of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
           Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-7"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>389
                     /note="HERV-W LTR like element"

  Query Match             4.3%;  Score 320.3;  DB 1;  Length 389;
  Best Local Similarity   90.5%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   15;  Indels   17;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| |||||||| ||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGATGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||| ||||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              ||||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||| ||||||||||||||||||:| |||||||||| ||||||||||||||||||
Db        241 TTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGAT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||| |||||:||| |||||||||||||||||| ||||:||||||||
Db        301 TCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||:|||||
Db        360 CCATCTTGGAAGCGGCCTGCTACCATCTTG 389


RESULT 922
AB066650
LOCUS       AB066650                389 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HB-11.
ACCESSION   AB066650
VERSION     AB066650.1  GI:18181892
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="HB-11"
                     /mol_type="mRNA"
                     /isolate="HB-11"
                     /db_xref="taxon:9606"
     LTR             1. .389
                     /note="HERV-W LTR family"

  Query Match             4.2%;  Score 318.3;  DB 1;  Length 389;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  351;  Conservative    5;  Mismatches   17;  Indels   17;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| |||||||| ||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||| ||||||||||||||||||| |||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGNTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              ||||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||| ||||||||||||||||||:| |||||||||| ||||||||||||||||||
Db        241 TTCTCTTCCGTGACCCACGGCTTNTAATAGAGCTATAACACTCACCACATGGCCCAAGAT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||| |||||:||| |||||||||||||||||| ||||:||||||||
Db        301 TCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||:|||||
Db        360 CCATCTTGGAAGTGGCCTGCTACCATCTTG 389


RESULT 923
AB066651
LOCUS       AB066651                389 bp    mRNA    linear   PRI 16-JAN-2002
DEFINITION  Homo sapiens mRNA, HERV-W LTR family, isolate:HB-12.
ACCESSION   AB066651
VERSION     AB066651.1  GI:18181893
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H.
  TITLE     Isolation of HERV-W LTR family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812964)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HB-12"
                     /db_xref="taxon:9606"
     LTR             1. .389
                     /note="HERV-W LTR family"

  Query Match             4.2%;  Score 317.7;  DB 1;  Length 389;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  351;  Conservative    5;  Mismatches   17;  Indels   17;  Gaps    2;
```

```
Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
              ||||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
              ||||||||||:|||||| ||||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
              |||||||||||||| ||||| || |||||||||||||||||| ||| ||||||||||||
Db        181 CACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCATTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
              |||||||| |||||||||||||| |||||:| |||||||||| ||||||||||||||| |
Db        241 TTCTCTTCCGTGACCCACGGCTTNTAATAGAGCTATAACACTCACCACATGGCCCAAGGT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
              |||||||||||||||| |||||:||| |||||||||||||||||| ||||:||||||||
Db        301 TCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||:|||||
Db        360 CCATCTTGGAAGCGGCCTGCTACCATCTTG 389
```

```
RESULT 924
AB079283
LOCUS       AB079283                372 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pongo pygmaeus DNA, HERV-W LTR element, isolate:OR-1.
ACCESSION   AB079283
VERSION     AB079283.1  GI:28201114
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="OR-1"
                     /db_xref="taxon:9600"
                     /note="HERV-W LTR element"
     LTR             <1. .>372
                     /note="HERV-W LTR like element"

  Query Match             4.2%;  Score 316.5;  DB 1;  Length 372;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  339;  Conservative    5;  Mismatches   28;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||| || |||||||||||||||||||||||||||| |||||||||
Db          1 CGAGCTGAGCTTTTGCTCGCCATCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||| |||||||||||||||| |||||||||||||||||| |:||| |
Db         61 TGACTCCCATCCCTCCGGATCCTGCAGGGTGTCCACTGTGCTCCTGATCCAGCAAGGCAC 120
```

```
Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              ||| ||| |||||| || ||||||||||||||||||||| |||||| |||||||||||| |
Db        121 CCACTGCTGCTCCCGATCGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCCG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||  | |||||||| |||||||||| |||||||||| |||||||||||||| |||||||
Db        181 GGTTCATCCTAATTGAACTGAACACTAGTCACTGGGTTTCATGGTTCTCTTCCGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| |||||||||| ||| ||||||||||||||||||||||||||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||||| ||||:|||||||||||||||||||| ||||
Db        301 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAACCGGC 359

Qy        508 CTGCTACCRTCTT 520
              |||||||||:||||
Db        360 CTGCTACCATCTT 372
```

```
RESULT 925
AB055870
LOCUS       AB055870                376 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:A549-1.
ACCESSION   AB055870
VERSION     AB055870.1  GI:15207779
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 376)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="A549-1"
                     /db_xref="taxon:9606"
                     /cell_line="A549"
     LTR             1. .376
                     /note="HERV-W family"

  Query Match             4.2%;  Score 316.4;  DB 1;  Length 376;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches  359;  Conservative    5;  Mismatches    9;  Indels    4;  Gaps    4;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          1 CGAGATGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCAGACCTGCCGT 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGC-GARGCG 266
              |||||||||||||| |||||||||||||||||||||||||||||||||||| ||:|||
Db         61 TGACTCCCATCCCTATGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCAGAGGCG 120

Qy        267 CCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCT 326
              |||||||||||||||||||| |||||||||||||||| ||||||||| |||||||||
Db        121 CCCATTGCCGCTCCCAATTGGGATAAAGGCTTGCCATTGTTCCTGCACGGGTAAGTGCCT 180

Qy        327 GGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACC 386
```

```
                 |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GGGTTTGTTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACC 240

Qy        387 CACGGCTTC-TAATAKAAC-TATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGA 444
                 ||||||||| |||||:||| ||||||||||||||||||||||||||||||||||| |||
Db        241 CACGGCTTCATAATAGAACTTATAACACTTACCACATGGCCCAAGATTCCATTCCTCGGA 300

Qy        445 ATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGC 504
                 ||||||||||:||| |||||||||||||||||||||:||||||||||||||||||||
Db        301 ATCCGTGAGGCCAA-GAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGC 359

Qy        505 GGCCTGCTACCRTCTTG 521
                 ||||||||||||:|||||
Db        360 GGCCTGCTACCATCTTG 376


RESULT 926
AB079281
LOCUS       AB079281                389 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-6.
ACCESSION   AB079281
VERSION     AB079281.1  GI:28201112
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-6"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>389
                     /note="HERV-W LTR like element"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 389;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  350;  Conservative    5;  Mismatches   18;  Indels   17;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCT----------------GTTTGCCACC 191
                 |||||||||||||||||||| ||||||||||||                ||||||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
                 |||||| ||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTG 311
                 ||||||||||:||||||| |||||||||||||||||||||||||||||||||| ||||
Db        121 TGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCTG 180

Qy        312 CACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGG 371
                 |||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        181 CACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGG 240

Qy        372 TTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGAT 431
                 |||||||| ||||||||||||||||||:| |||||||||| ||||||||||||||| |
Db        241 TTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGGT 300

Qy        432 TCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCA 491
                 |||||||||||||||| ||||||:||| |||| |||||||||||||| ||||:||||||||
Db        301 TCCATTCCTTGGAATCTGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCA 359

Qy        492 CCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
                 |||||||||||||||||||| | |:|||||
Db        360 CCATCTTGGAAGCGGCCTGCCAACATCTTG 389


RESULT 927
CQ053375/c
LOCUS       CQ053375                414 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 4195 from Patent WO0157270.
ACCESSION   CQ053375
VERSION     CQ053375.1  GI:41027881
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 4195 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN HBL100, SIGNAL = 1.9"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
                 ||| | ||||||||||||||||||||||||||||||||||||||||| ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
                 |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
                 ||||||| |||||||||||||||| | |||||| |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
                 |||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
                 ||||||| ||  |||||||| |||||  ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
                 |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
                 |||||| |||||  |||||||||||| ||||||   ||| |
Db          55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 928
CQ068513/c
LOCUS       CQ068513                414 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 4313 from Patent WO0157278.
ACCESSION   CQ068513
```

```
VERSION     CQ068513.1  GI:41038382
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 4313 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN HELA, SIGNAL = 1.6"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
              |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
              ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
              ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
              |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db          55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 929
CQ095583/c
LOCUS       CQ095583                414 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4442 from Patent WO0157272.
ACCESSION   CQ095583
VERSION     CQ095583.1  GI:41064609
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 4442 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN PLACENTA, SIGNAL = 6.2"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
              |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
              ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
              ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
              |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db          55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 930
CQ134328/c
LOCUS       CQ134328                414 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4350 from Patent WO0157276.
ACCESSION   CQ134328
VERSION     CQ134328.1  GI:41091684
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 4350 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.5"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | ||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
```

```
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| ||||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :|||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| ||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 931
CQ172831/c
LOCUS       CQ172831                414 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4227 from Patent WO0157274.
ACCESSION   CQ172831
VERSION     CQ172831.1  GI:41167567
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 4227 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN HEART, SIGNAL = 1.4"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | |||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :|||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| ||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 932
CQ202022/c
LOCUS       CQ202022                414 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4299 from Patent WO0157271.
ACCESSION   CQ202022
VERSION     CQ202022.1  GI:41187686
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 4299 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN BT474, SIGNAL = 1.5"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
              ||| | |||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
              ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :|||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| ||||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 933
```

```
CQ217589/c
LOCUS       CQ217589                414 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 4428 from Patent WO0157273.
ACCESSION   CQ217589
VERSION     CQ217589.1  GI:41198913
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 4428 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.2"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 934
CQ256153/c
LOCUS       CQ256153                414 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4414 from Patent WO0157277.
ACCESSION   CQ256153
VERSION     CQ256153.1  GI:41228633
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 4414 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.2"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| |||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| ||||||  |||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 935
CQ293239/c
LOCUS       CQ293239                414 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4344 from Patent WO0186003.
ACCESSION   CQ293239
VERSION     CQ293239.1  GI:41253816
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 4344 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN LUNG, SIGNAL = 1.7"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             ||||:|| ||||| ||| ||||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| |||||  ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 936
CQ330194/c
LOCUS       CQ330194                414 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 4288 from Patent WO0157275.
ACCESSION   CQ330194
VERSION     CQ330194.1  GI:41278761
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 4288 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN BRAIN, SIGNAL = 1.7"

  Query Match             4.2%;  Score 315.5;  DB 1;  Length 414;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  347;  Conservative    9;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       5396 AACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTG 5455
             ||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||| ||||
Db        414 AACCGTGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTG 355

Qy       5456 GATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCT 5515
             |||||| ||||| ||| ||||||||||||||||||||| ||||||||||||||||||||||
Db        354 GATACTACCAAACGAAACTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCT 295

Qy       5516 AGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCATAAAT 5575
             ||||||| |||||||||||||||| | ||||||  |||||||||||||| |||| || ||
Db        294 AGAGGATCTGCGCCTGCTCTTCAAGCGACAACCGTGAGGAAAGTAACTAGAATCGTAGAT 235

Qy       5576 CCCCCATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTC 5635
             ||||||||:||| |||| |||||||||||| |:||:||||:| |:|||| :||||||||
Db        234 -CCCCATGGCCCTACCTTGTCATATTTTTCTTTTTACTGTTCTCTTACCCCCTTTCACTC 176

Qy       5636 TCACTGCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTAT 5695
             ||||||| ||  |||||||| |||||  ||||||||||||||:|||:|||| ||||||
Db        175 TCACTGCTCCTTCTCCATGCTGCTGTACTACCAGTAGCTCCCCTTACCAAGAGCTTCTAT 116

Qy       5696 GGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAA 5755
             |||||||| ||| |||| |||||||||||||||||||||||| || |||:||| ||||
Db        115 GGAGAATGGAGCTTCCCAGAAATATTGATGCCCCATCGTATAGGGGTTTTTCTAAAGGAA 56

Qy       5756 CCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC 5796
              |||||| |||||  |||||||||| ||||||   ||| |
Db         55 ACCCCACTTTCACCACCCACACCCATGTGCCCCTGCACTTC 15


RESULT 937
CQ750358
LOCUS       CQ750358                357 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 36292 from Patent WO02068579.
ACCESSION   CQ750358
VERSION     CQ750358.1  GI:42381369
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 36292 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             4.2%;  Score 312.8;  DB 1;  Length 357;
  Best Local Similarity   92.1%;  Pred. No. 0;
  Matches  326;  Conservative    3;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       5199 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 5258
             |||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||||
Db          4 GCCAACTTAGTTGCAGACGTCACCTGCTTAGCCAAATATCAACAAGTTCTTAAAACATTA 63

Qy       5259 CAAGGAACCTATCCCTGAGAAGAGGGAAAAGAACTATTCCACCCWWGTGACATGGTATTA 5318
             |||||| ||| |||| |||||||||||||| ||| ||||||||||: |||||||| |||||
Db         64 CAAGGAGCCTGTCCCCGAGAAGAGGGAAAGGAATTATTCCACCCTGGTGACATGCTATTA 123

Qy       5319 GTCAAGTCCCTTCYCTCTAATTCCCCATCCCTAGATACATCCTGGGAAGGACCCTACCCA 5378
             | |||||||||||:|| ||||||||||||||||||||||||||||||| ||||||||||||
Db        124 GCCAAGTCCCTTCCCTTTAATTCCCCATCCCTAGATACATCCTGGGATGGACCCTACCCA 183

Qy       5379 GTCATTTTATYTACCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCAC 5438
             |||||||||:|||||||| ||||||||||||||||||||||||||||||||||||| ||||
Db        184 GTCATTTTATCTACCCCAACCTCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACCTCAC 243

Qy       5439 ACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCTAGC 5498
             ||| ||| |||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        244 ACTCGAGCCAAATCCTGGATACTGCCAAAGGAACCGAAAAATCCAGGAGACAACGCTAGC 303

Qy       5499 TATTCCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGA 5552
```

```
                ||||||||||||| |||||||||| || |||||||||| | ||||||||| ||
Db        304 TATTCCTGTGAACATCTAGAGGATCTGTGCCTGCTCTTCGAGCAACAACCATGA 357


RESULT 938
AB079276
LOCUS       AB079276                390 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-1.
ACCESSION   AB079276
VERSION     AB079276.1  GI:28201107
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 390)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .390
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-1"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>390
                     /note="HERV-W LTR like element"

  Query Match             4.1%;  Score 309.4;  DB 1;  Length 390;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  352;  Conservative    5;  Mismatches   16;  Indels   18;  Gaps    3;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG----------------TTTGCCACC 191
              |||| |||||||||||||||| |||||||||||||                 |||||||||
Db          1 CGAGNTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCACC 60

Qy        192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCC 251
              |||||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         61 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTCC 120

Qy        252 TGATCCAGCGARGCGC-CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCT 310
              ||||||||||:||||  || |||||||||||||||||||||||||||||||||| |||
Db        121 TGATCCAGCGAGGCGCNNCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCT 180

Qy        311 GCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATG 370
              |||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        181 GCACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATG 240

Qy        371 GTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGA 430
              |||||||| ||||||||||||||||||:| |||||||||| ||||||||||||||||||
Db        241 GTTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGA 300

Qy        431 TTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCC 490
              |||||||||||||||| |||||:||| |||||||||||||||| ||||:|||||||||
Db        301 TTCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCC 359

Qy        491 ACCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||||:|||||
Db        360 ACCATCTTGGAAGCGGCCTGCTACCATCTTG 390


RESULT 939
AB079280
LOCUS       AB079280                392 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-5.
ACCESSION   AB079280
VERSION     AB079280.1  GI:28201111
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 392)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .392
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-5"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>392
                     /note="HERV-W LTR like element"

  Query Match             4.1%;  Score 309.4;  DB 1;  Length 392;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  353;  Conservative    5;  Mismatches   15;  Indels   20;  Gaps    3;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG----------------TTTGCCA 189
              |||||||||||||||||||| |||||||||||||                 |||||||
Db          1 CGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTTGCCGCCACCACTGTTTTGCCA 60

Qy        190 CCACCGCANACCTGCC-GCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGC 248
              |||||||| ||||||| ||||||||||||| |||||||||||||||||||| ||||||
Db         61 CCACCGCAGACCTGCCGGCTGACTCCCATCCNTCTGGATCCTGCAGGGTGTCCACTGTGC 120

Qy        249 TCCTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNC 308
              |||||||||||||:||||||| |||||||||||||||||||||||||||||||||||| |
Db        121 TCCTGATCCAGCGAGGCGCCCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTC 180

Qy        309 CTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCA 368
              |||||||||||||||| ||||| || |||||||||||||||||| ||||||||||||
Db        181 CTGCACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCA 240

Qy        369 TGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAA 428
              |||||||||| ||||||||||||||||||:| |||||||||| ||||||||||||||
Db        241 TGGTTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAA 300

Qy        429 GATTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTG 488
              |||||||||||||||||||| |||||:||| |||||||||||||||| ||||:|||||
Db        301 GATTCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTG 359

Qy        489 CCACCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||||||||||||||||||||||||:|||||
Db        360 CCACCATCTTGGAAGCGGCCTGCTACCATCTTG 392


RESULT 940
AB079278
LOCUS       AB079278                389 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-3.
ACCESSION   AB079278
VERSION     AB079278.1  GI:28201109
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 389)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-3"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>389
                     /note="HERV-W LTR like element"

  Query Match             4.1%;  Score 308.2;  DB 1;  Length 389;
  Best Local Similarity   90.0%;  Pred. No. 0;
  Matches  352;  Conservative    5;  Mismatches   16;  Indels   18;  Gaps    3;

Qy        147 TCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC----------------TGTTTGCCAC 190
              ||||||||||||||||||||||| ||||||||||                ||||||||||
Db          1 TCGAGCTGAGCTTTTGCTCACCATCCACCACTGCTGTTTGCCACCACCACTGTTTGCCAC 60

Qy        191 CACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTC 250
              ||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         61 CACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCACTGTGCTC 120

Qy        251 CTGATCCAGCGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCT 310
              ||||||||||||:|||  || |||||||||||||||||||||||||||||||||| |||
Db        121 CTGATCCAGCGAGGCGAGCACTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTTCCT 180

Qy        311 GCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATG 370
              |||||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        181 GCACGGCTAAGTGCCCGGGTTCGTCCTAATTGAGCTGAACACTAGTCACTGGGTTCCATG 240

Qy        371 GTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGA 430
              ||||||||| |||||||||||||||||||:| |||||||||| ||||||||||||||||
Db        241 GTTCTCTTCCGTGACCCACGGCTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGA 300

Qy        431 TTCCATTCCTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCC 490
              |||||||||||||||||| |||||:||| |||||||||||||||||| ||||:|||||||
Db        301 TTCCATTCCTTGGAATCCATGAGGCCAA-GAACTCCAGGTCAGAGAACACGAGGCTTGCC 359

Qy        491 ACCATCTTGGAAGCGGCCTGCTACCRTCTTG 521
              |||||| ||||||||||||||||||:|||||
Db        360 ACCATC-TGGAAGCGGCCTGCTACCATCTTG 389


RESULT 941
BV190586/c
LOCUS       BV190586                401 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm167639 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV190586
VERSION     BV190586.1  GI:48032100
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 401.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>401

  Query Match             4.1%;  Score 307.6;  DB 1;  Length 401;
  Best Local Similarity   92.6%;  Pred. No. 0;
  Matches  325;  Conservative    7;  Mismatches   17;  Indels    2;  Gaps    1;

Qy       7232 GGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYT 7291
              |||||||| |||||| |||||||||||||||||||||||||||||| ||||||| ||:|
Db        401 GGTTTTCCTATTGAGAGGGGGGACTGAGAGACAGGACTAGCTGGATTCCCTAGGCCGACT 342

Qy       7292 AAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCA 7351
              ||||||||:|||||||||:|||||||||||||  ||||||||||||||||||||||||||
Db        341 AAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTTAAACACGGGGCTTGCA 282

Qy       7352 ACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGG 7411
              ||||||:|||||||  |||||||  ||||||||||||||||||||||||||||| |||||
Db        281 ACTTAGCTCACACCCAACCAATC--AGAGCTCACTAAAATGCTAATTAGGCAAAAACAGG 224

Qy       7412 AGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGGACAATGATCGG 7471
              |||||||||||||||||||||||||||||:|||||||||||||||||| ||| ||||||
Db        223 AGGTAAAGAAATAGCCAATCATYTATTGCCTGAGAGCACAGCAGGAGGGGCAAGGATCGG 164

Qy       7472 GATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGT 7531
              |||||||||||| |  ||||||||||||||||||||||||||||||||||||||||||
Db        163 GATATAAACCCAGGCATTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGT 104

Qy       7532 ATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||||  |||||||||||||||||||||||||:|| |:
Db        103 ATGGGAGCTCTGTTTTCACTCTATTTCACTCTATTAAATCTTGCAACTTCA 53


RESULT 942
BD003152
LOCUS       BD003152                374 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Secretory proteins and polynucleotides encoding the same.
ACCESSION   BD003152
VERSION     BD003152.1  GI:18631113
KEYWORDS    JP 2001501813-A/32.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Mccoy,J.M., Levallie,E.R., Racie,L.A., Merberg,D., Treacy,M.,
            Evans,C. and Spaulding,V.
  TITLE     Secretory proteins and polynucleotides encoding the same
  JOURNAL   Patent: JP 2001501813-A 32 13-FEB-2001;
            GENETICS INSTITUTE INC,KENNETH JACOBS
COMMENT     OS   Unidentified
            PN   JP 2001501813-A/32
            PD   13-FEB-2001
            PF   23-JUL-1997 JP 1998508939
            PR   26-JUL-1996 US    08/686878
            PI   JOHN M MCCOY,EDWARD R LEVALLIE,LISA A RACIE,DAVID MERBERG, PI
             MAURICE TREACY,
```

```
            PI   CHERYL EVANS,VIKKI SPAULDING
            PC   C12N15/09,A61K38/00,A61K48/00,A61P3/02,A61P7/02,A61P7/08, PC
            A61P35/00,
            PC   A61P37/00,C07K14/47,C12N5/10,C12P21/02,C12N15/00,C12N5/00, PC
             A61K37/02
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .374
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:||||||||||||||||||:||||||:||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             |  |||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             |  |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 943
I80071
LOCUS       I80071                   374 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 47 from patent US 5708157.
ACCESSION   I80071
VERSION     I80071.1  GI:3208361
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Evans,C. and Spaulding,V.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5708157-A 47 13-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             4.1%;  Score 306.3;  DB 1;  Length 374;
  Best Local Similarity   91.0%;  Pred. No. 0;
  Matches  315;  Conservative    6;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       4658 CTGCCTYATCGCCAAGCTCCTTCAGGARAACAAARAACAGGCCATTACCCTGRARAARAC 4717
             ||||||:||||||||||||||||||:||||||:||||||||||||||||:|:|| ||
Db         29 CTGCCTTATCGCCAAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAAAANAC 88

Qy       4718 TGGCAACTGATTTTACCCACAAGCCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG 4777
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||||
Db         89 TGGCAACTGATTTTACCCACAAGCCCNAACCTCAGGGATTTCAGTATCTACTANTCTGGG 148

Qy       4778 TARATACTTTCACGGGTTGGGCARAGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGG 4837
             |  |||||||||||||||||||:| |||||||||||| |||||||||| |||| ||| |
Db        149 TNNATACTTTCACGGGTTGGGCAGAAGCCTTCCCCTGTTGGACAGAAAANGCCCNAGANG 208

Qy       4838 TAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGA 4897
             |  |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        209 TTNTAAAGGCACTAGTTCATGAAATAATTCCCANATTCGGACTTCCCCGAGGCTTACAGA 268

Qy       4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCA-GGGAGTATCCCAGGCGTTAGGT 4956
              |||| ||  |||||||||||||||||||||| ||||| |||||| |||| |||||| |||
Db        269 NTGACNATNNCCCTGCTTTCCAGGCCACAGTNACCCAGGGGAGTNTCCCNGGCGTTNGGT 328

Qy       4957 ATACGATATCACTTACACTGCGCCTGAAGGCCACAGTCCTCAGGGA 5002
              |||||||||||||||||||||||||||| |||||||||||| ||||
Db        329 NTACGATATCACTTACACTGCGCCTGAANGCCACAGTCCTCNGGGA 374


RESULT 944
AB071970
LOCUS       AB071970                 375 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch1.
ACCESSION   AB071970
VERSION     AB071970.1  GI:19698205
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 375)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch1"
                     /db_xref="taxon:9606"
     LTR             1. .375
                     /note="HERV-W LTR"

  Query Match             4.0%;  Score 300.1;  DB 1;  Length 375;
  Best Local Similarity   89.4%;  Pred. No. 0;
  Matches  336;  Conservative    5;  Mismatches   32;  Indels    3;  Gaps    2;

Qy       148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
             |||||||||||||| |||| || |||||||||||||||||| ||  |||| |||||| ||
Db         1 CGAGCTGAGCTTTGGCTCGCCATCCACCACTGCTGTTTGCCGCCGTCGCACACCTGCTGC 60

Qy       208 TGACTCCCATCCCTCTGGATCCTGCAGG--GTGTCCGCTGTGCTCCTGATCCAGCGARGC 265
             |||||||||||||| |||||| |||||  |||||||||||||||||||||||||||:|
Db        61 TGACTCCCATCCCTCCGGATCCAGCAGGGTGTGTCCGCTGTGCTCCTGATCCAGCGAGGT 120

Qy       266 GCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCC 325
             |||||||||||||||  ||||| |||||||||||||||||| ||||||||||||||||||
Db        121 GCCCATTGCCGCTCCTGATTGGACTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTGCC 180

Qy       326 TGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGAC 385
             |||||  ||  |||||  |||||||||||| |||||||||||| ||||||||||  ||||
Db        181 CGGGTTCGTCCTAATCCAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCTTGAC 240
```

```
Qy        386 CCACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAA 445
              ||||||||||||||:| |||||||||| ||| ||||||||||||||||||||||||||
Db        241 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAA 300

Qy        446 TCCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCG 505
              || |||||:||| |||| |||||||||||| ||||:|||||||||||||||||||||||
Db        301 TCTGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCG 359

Qy        506 GCCTGCTACCRTCTTG 521
              |||||| | |:|||||
Db        360 GCCTGCCAACATCTTG 375


RESULT 945
AB079294
LOCUS       AB079294                372 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Hylobates agilis DNA, HERV-W LTR element, isolate:GI-4.
ACCESSION   AB079294
VERSION     AB079294.1  GI:28201125
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GI-4"
                     /db_xref="taxon:9579"
                     /note="HERV-W LTR element"
     LTR             <1. .>372
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 295.7;  DB 1;  Length 372;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches  326;  Conservative    5;  Mismatches   41;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||| | |||||| |||||||||||  ||| |||| ||  ||| |
Db          1 CGAGCTGAGCTTTTGCTCCCTGTCCACTACTGCTGTTTGCAGCCATCGCAGACTCGCCAC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||| |||||||||  ||||| |||||||||||| |||||||||||||||||| ||:||||
Db         61 TGACTTCCATCCCTCCAGATCCAGCAGGGTGTCCACTGTGCTCCTGATCCAGTGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||| |||||| ||  |||||||||||||||||||| |||||| ||||||||||||||
Db        121 CCATTGCCACTCCCCATCTGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||  || ||||| |||||||||||||| ||||| ||||||| |||||| ||||||||||||
Db        181 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTCGGTTCCACGGTTCTTTTCTGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||| |||||||||:| |||||||||| ||||||||||||||||||||||||||||||||||
Db        241 ACGACTTCTAATAGAGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||| || |||||||||| ||||:|||||||||||| |||||||| |||
Db        301 TGTGAGGCCAA-GAACCCCTGGTCAGAGAACACGAGGCTTGCCACCAACTTGGAAGTGGC 359

Qy        508 CTGCTACCRTCTT 520
              |||| ||:| ||
Db        360 CTGCCGCCATTTT 372


RESULT 946
AB079274
LOCUS       AB079274                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-6.
ACCESSION   AB079274
VERSION     AB079274.1  GI:28201105
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-6"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 295.1;  DB 1;  Length 373;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  326;  Conservative    5;  Mismatches   42;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||| |||| |  |||||||||||||||||||| |   |||| |||||||||
Db          1 CGAGCTGAGCTTTAGCTCGCTTTCCACCACTGCTGTTTGCCGCTGTCGCAGACCTGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||| |||||| |||||||||||||||||| |||||| ||||| |:|| |
Db         61 TGACTCCCATCCCTCCGGATCCGGCAGGGTGTCCGCTGTGTTCCTGAACCAGCAAGGCAC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||  |  ||||||||||||||||||||||| |||||| ||||||||||||
Db        121 CCATTGCCGCTTCTGATTGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCCA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || ||||| |||||||||||||| |||||||||||| |||||| |||  ||||||
Db        181 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTTTTCCATGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              | ||||||||||||:| |||||||||| || |||||||||||||||||||||||||||||||
Db        241 ATGGCTTCTAATAGAGCTATAACACTCACCGCACGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||| |||||||||||||| ||||:||| ||||||||||||||||||||
Db        301 CGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTAGCCACCATCTTGGAAGCGGC 359

Qy        508 CTGCTACCRTCTTG 521
              |  | |||:||| |
Db        360 CCACCACCATCTGG 373


RESULT 947
```

```
BV195837/c
LOCUS       BV195837                401 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm186014 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV195837
VERSION     BV195837.1  GI:48039806
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 401.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>401

  Query Match             3.9%;  Score 293.8;  DB 1;  Length 401;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches  329;  Conservative    8;  Mismatches   40;  Indels   12;  Gaps    1;

Qy       7013 CCAGAATCGAAGCTGTRAAACTA------------CAAATGGAGCCCAAGATGCAGTCCA 7060
              ||||||||||||||||:|| |||            |||||||||||| ||||||| ||||
Db        401 CCAGAATCGAAGCTGTAAACCTAAAAATTGTTCTTCAAATGGAGCCCCAGATGCAATCCA 342

Qy       7061 AGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCCCACGATCTGATGTTAATG 7120
              | |||||||||||||||||||| |||||| ||||||:||||||| | |  ||||||||||
Db        341 TGGCTAAGATCTACCGCAGACCTCTGGACTGGCCTGCTAGCCCATGCTTCGATGTTAATG 282

Qy       7121 ACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATT 7180
              ||||  |||||||||||| ||| ||||||||| |||||||||| ||| || |||||||||
Db        281 ACATTGAAGGCACCCCTCTTGAAGAAATCTCAACTGCACAACCCCTAGTATGCCCCAATT 222

Qy       7181 CAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCT 7240
              ||||||||||||||||||||:||: || |||||||||||||||||||| |||||||||
Db        221 CAGCAGGAAGCAGTTAGAGCRGTCATCAGCCAACCTCCCCAACAGCACTTGGGTTTTCCT 162

Qy       7241 GTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCY 7300
              ||||||| |||| |||||||||||||||| ||||||||||||||| ||:|||||||||:
Db        161 GTTGAGAAGGGGTACTGAGAGACAGGACTGGCTGGATTTCCTAGGCCGACTAAGAATCCT 102

Qy       7301 TAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTC 7360
              |||||||||:|||||| |||    || ||||||||||||||||||||||||||:||
Db        101 TAAGCCTAGCTGGGAAGTTGAGTGTATGCACCTTTAAACACGGGGCTTGCAACTTAGCTC 42

Qy       7361 ACACCTGACCAATCAGAGAGCTCACTAAA 7389
              |||||  |||||||||  ||   | ||||
Db         41 ACACCCAACCAATCAGGTAGTAAAGTAAA 13


RESULT 948
AB071974
LOCUS       AB071974                373 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch5.
ACCESSION   AB071974
VERSION     AB071974.1  GI:19698209
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch5"
                     /db_xref="taxon:9606"
     LTR             1. .373
                     /note="HERV-W LTR"

  Query Match             3.9%;  Score 293.5;  DB 1;  Length 373;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   43;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||| |||| |  ||||||||||||||||||| |   |||| ||||||| |
Db          1 CGAGCTGAGCTTTCGCTCGCTTTCCACCACTGCTGTTTGCCGCTGTCGCAGACCTGCCAC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||| |||||| ||||| |||||||||||||||||| |||||||:|| |
Db         61 TGACTCCCATCCCTCCGGATCCGGCAGGCTGTCCGCTGTGCTCCTGAACCAGCGAGGCAC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              ||||||||  |||  |||||||||||||||||| |||| ||||||||||||||||||||
Db        121 CCATTGCCAATCCTGATTGGGCTAAAGGCTTGCCGTTGTTCCTGCACGGCTAAGTGCCCA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||  | |||||||||||||||||| |||||||||||| ||||||||||  ||||||
Db        181 GGTTCATCCTAATTGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCATGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              | ||||||||||:| |||||||||| ||| || ||||||||||||||||||||||||||
Db        241 ATGGCTTCTAATAGAGCTATAACACTCACCGCACGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||| ||||:||| |||||||||||||| |||
Db        301 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTAGCCACCATCTTGGAAGTGGC 359

Qy        508 CTGCTACCRTCTTG 521
              || | |||:||| |
Db        360 CTACCACCATCTGG 373


RESULT 949
CQ114106/c
LOCUS       CQ114106                336 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 22965 from Patent WO0157272.
ACCESSION   CQ114106
VERSION     CQ114106.1  GI:41083976
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 22965 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN PLACENTA, SIGNAL = 2.6
                     SWISSPROT HIT: P26808, EVALUE 6.00e-18
                     NT HIT: AF009668.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||: ||||||||||:||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: ||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:||||||:||||||||||||: ||| ||||||||||||||||||||||||||
Db         96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 950
CQ152987/c
LOCUS       CQ152987                336 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 23009 from Patent WO0157276.
ACCESSION   CQ152987
VERSION     CQ152987.1  GI:41160337
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 23009 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.5
                     SWISSPROT HIT: P26808, EVALUE 6.00e-18
                     NT HIT: AF009668.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||: ||||||||||:||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: ||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:||||||:||||||||||||: ||| ||||||||||||||||||||||||||
Db         96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 951
CQ186062/c
LOCUS       CQ186062                336 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17458 from Patent WO0157274.
ACCESSION   CQ186062
VERSION     CQ186062.1  GI:41181077
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 17458 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN HEART, SIGNAL = 1.3
                     SWISSPROT HIT: P26808, EVALUE 6.00e-18
                     NT HIT: AF009668.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||: ||||||||||:||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217
```

```
Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 952
CQ236297/c
LOCUS       CQ236297                336 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 23136 from Patent WO0157273.
ACCESSION   CQ236297
VERSION     CQ236297.1  GI:41219575
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 23136 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .336
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC022555.2
                      EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
                      SWISSPROT HIT: P26808, EVALUE 6.00e-18
                      NT HIT: AF009668.1, EVALUE 0.00e+00
                      EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||  |||||||||||||||||:||||::|||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97
```

```
Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 953
CQ273879/c
LOCUS       CQ273879                336 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 22140 from Patent WO0157277.
ACCESSION   CQ273879
VERSION     CQ273879.1  GI:41246483
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 22140 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .336
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC022555.2
                      EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
                      SWISSPROT HIT: P26808, EVALUE 6.00e-18
                      NT HIT: AF009668.1, EVALUE 0.00e+00
                      EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||  |||||||||||||||||:||||::|||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:||: |||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:|||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             |||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 954
CQ311149/c
LOCUS       CQ311149                336 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 22254 from Patent WO0186003.
ACCESSION   CQ311149
VERSION     CQ311149.1  GI:41271726
```

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 22254 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN LUNG, SIGNAL = 2
                     SWISSPROT HIT: P26808, EVALUE 6.00e-18
                     NT HIT: AF009668.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:|| ||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 955
CQ348183/c
LOCUS       CQ348183                336 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 22277 from Patent WO0157275.
ACCESSION   CQ348183
VERSION     CQ348183.1  GI:41297254
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 22277 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC022555.2
                     EXPRESSED IN BRAIN, SIGNAL = 1.5
                     SWISSPROT HIT: P26808, EVALUE 6.00e-18
                     NT HIT: AF009668.1, EVALUE 0.00e+00
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-130"

  Query Match             3.9%;  Score 292.1;  DB 1;  Length 336;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches  302;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3628 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAAAGGA 3687
             ||:||||||||||  ||||||||||||||||||:|||||::||||||||||||||||||||
Db        336 GGAAGAGATCTTATCGTGTGGACATCTCATGATGTGAACGGCATACTCACTGCTAAAGGA 277

Qy       3688 GACTTGTGGCTGTCAGACAACYGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCA 3747
             |||||||||||||||||||||: ||||||||||:|||||||||||||||||||:| |||
Db        276 GACTTGTGGCTGTCAGACAACCATTTACTTAAATATCAGGCTCTATTACTTGAAGGACCA 217

Qy       3748 GTGCTGCRACTGTGCACTTGTGCAACTCTTAACCCAGYCNCATTTCTTCCAGACAATGAA 3807
             |||||||:||||  ||| ||||||||||||||||||:| ||||||||||||||||||||
Db        216 GTGCTGCGACTGCACACCTGTGCAACTCTTAACCCAGTCACATTTCTTCCAGACAATGAA 157

Qy       3808 GAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCGAGGGGAC 3867
             ||||||| |:|| ||||||||||||:||||| |||||||||| || ||||||||||||
Db        156 GAAAAGACAGAACGTAACTGTCAACAAGTAATTGCTCAAACCTACGCTACTCGAGGGGAC 97

Qy       3868 CTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTT 3927
             |||:||||:|||||:||||||||||: ||| ||||||||||||||||||||||||||||
Db          96 CTTCTAGAGGTTCCCTTGACTGATCCCTACC-TCAACTTGTATACTGATGGAAGTTCCTT 38

Qy       3928 TGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964
             ||||||||||||||||||||| |: | ||||||||||
Db          37 TGTAGAAAAAGGACTTCGAAAGGCAGAGTATGCAGTG 1


RESULT 956
AB079291
LOCUS       AB079291                372 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Hylobates agilis DNA, HERV-W LTR element, isolate:GI-1.
ACCESSION   AB079291
VERSION     AB079291.1  GI:28201122
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GI-1"
                     /db_xref="taxon:9579"
                     /note="HERV-W LTR element"
     LTR             <1. .>372
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 291.5;  DB 1;  Length 372;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  323;  Conservative    5;  Mismatches   43;  Indels    1;  Gaps    1;
```

```
Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||  ||||||  | ||||||||||||||||  ||| | || |||  || |
Db          1 CGAGCTGAGCTTTCACTCACCATTCACCACTGCTGTTTGCTGCCATCACAGACCCCCCAC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||  |||||  |||||||||||||||||||||||||||||||| ||:||||
Db         61 TGACTCCCATCCCTCCGGATCTGGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||| ||  ||||||||||||||||||| ||||||  ||||||||||
Db        121 CCATTGCCGCTCCCCATCAGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCCA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||||||  |||||  |||||  |||||||  |||||  |||||||  |||||||||  |||||||
Db        181 GGTTTGTCCTAATGGAGCTCAACACTAGTCACTAGGTTCCACGGTTCTCTTACGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| ||||||||||  ||  |||||||||||||||||||||||| |||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACTGCATGGCCCAAGATTCCATTCCTTAGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              | |||||:||| |||| |||||| |||||| ||||:|||||||| ||||||||||| ||
Db        301 CCTGAGGCCAA-GAACCCCAGGTGAGAGAACACGAGGCTTGCCATCATCTTGGAAGCAGC 359

Qy        508 CTGCTACCRTCT 519
              |||| |||:|||
Db        360 CTGCCACCATCT 371


RESULT 957
AB079277
LOCUS       AB079277                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Gorilla gorilla DNA, HERV-W LTR element, isolate:GO-2.
ACCESSION   AB079277
VERSION     AB079277.1  GI:28201108
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="GO-2"
                     /db_xref="taxon:9593"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 291.5;  DB 1;  Length 373;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  323;  Conservative    5;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||  ||| |  |||||||||||||||||| |    |||| |||||| ||
Db          1 CGAGCTGAGCTTTCACTCGCTTTCCACCACTGCTGTTTGCCGCTGTCGCAGACCTGCTGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||| || ||||||  |||||||||||||||||||||||||||||||:|| |
Db         61 TGACTCCCATCCCTCCGGATCTGGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||  |||||| |||||||||||||||||| |||| ||||||||||| ||
Db        121 CCATTGCCGCTCCCCATCAGGCTAAAGGCTTGCCATTGTTCCTGCACGGCTAAGTACCCA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || ||||| |||||||||||||  |||||||||||  ||||||||||  ||||||
Db        181 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCATGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              | ||||||||||:| ||||||||| ||| || |||||||||||||||||||||||||||
Db        241 ATGGCTTCTAATAGAGCTATAACACTCACCGCACGGCCCAAGATTCCATTCCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              |||||||:||| |||| |||||||||||||| ||||:||| |||||||||||||||| ||
Db        301 CGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTAGCCACCATCTTGGAAGTGGT 359

Qy        508 CTGCTACCRTCTT 520
              | |||:||||
Db        360 TCACCACCATCTT 372


RESULT 958
AB079286
LOCUS       AB079286                372 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pongo pygmaeus DNA, HERV-W LTR element, isolate:OR-4.
ACCESSION   AB079286
VERSION     AB079286.1  GI:28201117
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="OR-4"
                     /db_xref="taxon:9600"
                     /note="HERV-W LTR element"
     LTR             <1. .>372
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 289.9;  DB 1;  Length 372;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  322;  Conservative    5;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||  |||||| | |||||||||| ||||  ||| | || |||  || |
Db          1 CGAGCTGAGCTTTCACTCACCATTCACCACTGCTGCTTGCTGCCATCACAGACCCCCCAC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||||  |||||  |||||||||||||||||||||||||||||||| ||:||||
Db         61 TGACTCCCATCCCTCCGGATCTGGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||| ||  |||||||||||||||||||| ||||||  ||||||||||
Db        121 CCATTGCCGCTCCCCATCAGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCCA 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
```

```
                |||||||| |||||| |||||| |||||||| |||||| |||||||| |||||||||  ||||||||
Db        181 GGTTTGTCCTAATGGAGCTCAACACTAGTCACTAGGTTCCACGGTTCTCTTACGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
                |||||||||||||:| |||||||||| ||  |||||||||||||||||||||||| |||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACTGCATGGCCCAAGATTCCATTCCTTAGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
                | |||||:||| |||| |||||| |||||| ||||:|||||||| |||||||||||| ||
Db        301 CCTGAGGCCAA-GAACCCCAGGTGAGAGAACACGAGGCTTGCCATCATCTTGGAAGCAGC 359

Qy        508 CTGCTACCRTCT 519
                |||| |||:|||
Db        360 CTGCCACCATCT 371


RESULT 959
AB079296
LOCUS       AB079296                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Hylobates agilis DNA, HERV-W LTR element, isolate:GI-6.
ACCESSION   AB079296
VERSION     AB079296.1  GI:28201127
KEYWORDS    .
SOURCE      Hylobates agilis (agile gibbon)
  ORGANISM  Hylobates agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hylobatidae; Hylobates.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Hylobates agilis"
                     /mol_type="genomic DNA"
                     /isolate="GI-6"
                     /db_xref="taxon:9579"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             3.9%;  Score 288.7;  DB 1;  Length 373;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  322;  Conservative    5;  Mismatches   46;  Indels    1;  Gaps    1;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
                |||||||||||||  |||||| |||||||||||||||| ||| |    ||| ||||||| |
Db          1 CGAGCTGAGCTTTCACTCACCATCCACCACTGCTGTTAGCCGCTGTGGCAGACCTGCCCC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
                ||||| |||||||||| |||||| |||||||||| ||||||||||||||||||||:||||
Db         61 TGACTTCCATCCCTCCGGATCCGGCAGGGTGTCTGCTGTGCTCCTGATCCAGCGAGGCGC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
                ||||| | |||||| || |||||||||||||||| ||| |||||| |||||||||| |
Db        121 CCATTACTGCTCCCGATCGGGCTAAAGGCTTGCCACTGTTCCTGCATGGCTAAGTGCCCG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
                |||| || |||||  |||||||| ||| |||||||||||||| |||||||||| |||||||
Db        181 GGTTCGTCCTAATCAAGCTGAACCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
                |||||||||||||:| |||||||||| ||  |||||||||||||||||||||||| |||||||||
Db        241 ACGGCTTCTAATAGAGCTATAACACTCACTGCATGGCCCAAGATTCCATTTCTTGGAATC 300

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
                |||||||:  | |||| |||||||| || | ||||:|||||||||||||||||||| |
Db        301 CGTGAGGCTGA-GAACCCCAGGTCACAGGACACGAGGCTTGCCACCATCTTGGAAGCAGT 359

Qy        508 CTGCTACCRTCTTG 521
                |||| |||:| |||
Db        360 CTGCCACCATTTTG 373


RESULT 960
AY208135
LOCUS       AY208135                329 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens isolate 4 endogenous retrovirus HERV-W long terminal
            repeat, partial sequence.
ACCESSION   AY208135
VERSION     AY208135.1  GI:28932686
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
   PUBMED   12867655
REFERENCE   2  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="4"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>329

  Query Match             3.9%;  Score 288.6;  DB 1;  Length 329;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches  294;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
                |||||||||||||||||||||||| ||:||| |||||:|||||||||:||||||||||||
Db          1 AGGACTAGCTGGATTTCCTAGGCCGACTAACAATCCCTAAGCCTAGCTGGGAAGGTGACC 60

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
                | |||||||||||||| ||||||||||||||||:|||||||||||||||||||||||||||
Db         61 GCTTCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 120

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
                |||||||||||||||||||| ||||||||||||||||||||||||:||||||:||
Db        121 ACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 180

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
                |||||||||| ||||||||||  |||||||||||||||| |  ||||||||||||||||
Db        181 AGAGCACAGCGGGAGGGACAAGCATCGGGATATAAACCCAGGCATTCGAGCCGGCAACGG 240

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
                |||||  || |||||||||||||||||||||||||| ||||||||||||  ||||||||||||
Db        241 CAACCACCGTTGGGTCCCCTCCCTTTGTATGGGGGCTCTGTTTTCACTCTATTTCACTCT 300

Qy       7564 ATTAAATCTTGCARCTGCR 7582
                |||||||||||||:||||:
Db        301 ATTAAATCTTGCAACTGCA 319


RESULT 961
AY208134
```

```
LOCUS       AY208134                329 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens isolate 3 endogenous retrovirus HERV-W long terminal
            repeat, partial sequence.
ACCESSION   AY208134
VERSION     AY208134.1  GI:28932685
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
   PUBMED   12867655
REFERENCE   2  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="3"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>329

  Query Match             3.8%;  Score 287;  DB 1;  Length 329;
  Best Local Similarity   91.8%;  Pred. No. 0;
  Matches  293;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             ||||||||||||||||||||||||||:|||||||||:|||||||||:||||||||||||
Db          1 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACG 60

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              ||||||||||||||||| ||||||||||||||:|||||||  |||||||  ||||||||
Db         61 GCATCCACCTTTAAACACAGGGCTTGCAACTTAGCTCACACCCAACCAATCGGAGAGCTC 120

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              |||||||||||||||||||| ||||||||||||||||||||||||:||||||:||
Db        121 ACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 180

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||||||| ||| |||||| |||||||||||||||| |  ||| |||| |||||||
Db        181 AGAGCACAGCGGGAAGGACAAGGATCGGGATATAAACCCAGGCATTCAAGCCAGCAACGG 240

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              |||||||||||||||||||||| ||||||||||||||||||  ||||||||||||
Db        241 CAACCCCCTTTGGGTCCCCTCCCCTTGTATGGGAGCTCTGTTTTCACTCTATTTCACTCT 300

Qy       7564 ATTAAATCTTGCARCTGCR 7582
              |||||||||||||:||||:
Db        301 ATTAAATCTTGCAACTGCA 319


RESULT 962
CQ081963/c
LOCUS       CQ081963                387 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 17763 from Patent WO0157278.
ACCESSION   CQ081963
VERSION     CQ081963.1  GI:41051832
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 17763 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004909.2
                     EXPRESSED IN HELA, SIGNAL = 1.5
                     EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
                     SWISSPROT HIT: P34109, EVALUE 2.50e-01
                     NT HIT: gi11419444, EVALUE 3.00e-96"

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| ||||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||| |||||| |||||||||   |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||| || |:|||||||||| | |||||||| ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 963
CQ116633/c
LOCUS       CQ116633                387 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25492 from Patent WO0157272.
ACCESSION   CQ116633
VERSION     CQ116633.1  GI:41086503
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 25492 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

/note="MAP TO AC004909.2
EXPRESSED IN PLACENTA, SIGNAL = 4.7
EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
SWISSPROT HIT: P34109, EVALUE 2.50e-01
NT HIT: gi|1419444, EVALUE 3.00e-96"

Query Match 3.8%; Score 285.7; DB 1; Length 387;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 325; Conservative 5; Mismatches 56; Indels 1; Gaps 1;

```
Qy      160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
            || | | ||| |||||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db      386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy      220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
            ||| |||| |||||||||| |||| |||||| |||||| |:| |||||||||| |||
Db      326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy      280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
            || || |||||||||||||||||||| |||||| |||||||||   |||| || || |
Db      266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy      340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
            |  ||||||| ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db      206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy      400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
            |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db      146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy      460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
            |||| |||||||||||| || |:|||||||| | |||||||| ||| ||| ||:| |
Db       87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy      520 TGGAAGTGGTTCACCACCATCTTGGGA 546
            |||||||||    ||||||||||||||
Db       27 TGGAAGTGGCCTGCCACCATCTTGGGA 1
```

RESULT 964
CQ238553/c
LOCUS CQ238553 387 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 25392 from Patent WO0157273.
ACCESSION CQ238553
VERSION CQ238553.1 GI:41221843
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000 (03.10.00)<150> US 60/236,359<151> 27 September 2000 (27.09.00)<150> US 60/234,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170> Molecular Dynamics Sequence Listing Engine
JOURNAL Patent: WO 0157273-A 25392 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .387
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC004909.2
EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
SWISSPROT HIT: P34109, EVALUE 2.50e-01
NT HIT: gi|1419444, EVALUE 3.00e-96"

Query Match 3.8%; Score 285.7; DB 1; Length 387;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 325; Conservative 5; Mismatches 56; Indels 1; Gaps 1;

```
Qy      160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
            || | | ||| |||||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db      386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy      220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
            ||| |||| |||||||||| |||| |||||| |||||| |:| |||||||||| |||
Db      326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy      280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
            || || |||||||||||||||||||| |||||| |||||||||   |||| || || |
Db      266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy      340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
            |  ||||||| ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db      206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy      400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
            |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db      146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy      460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
            |||| |||||||||||| || |:|||||||| | |||||||| ||| ||| ||:| |
Db       87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy      520 TGGAAGTGGTTCACCACCATCTTGGGA 546
            |||||||||    ||||||||||||||
Db       27 TGGAAGTGGCCTGCCACCATCTTGGGA 1
```

RESULT 965
CQ276189/c
LOCUS CQ276189 387 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 24450 from Patent WO0157277.
ACCESSION CQ276189
VERSION CQ276189.1 GI:41248793
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 24450 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .387
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC004909.2
EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
SWISSPROT HIT: P34109, EVALUE 2.50e-01
NT HIT: gi|1419444, EVALUE 3.00e-96"

Query Match 3.8%; Score 285.7; DB 1; Length 387;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 325; Conservative 5; Mismatches 56; Indels 1; Gaps 1;

```
Qy      160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
            || | | ||| |||||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db      386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy      220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
```

```
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| ||||||||||    |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | ||||||||  ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 966
CQ313183/c
LOCUS       CQ313183                387 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24288 from Patent WO0186003.
ACCESSION   CQ313183
VERSION     CQ313183.1  GI:41273760
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 24288 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004909.2
                     EXPRESSED IN LUNG, SIGNAL = 1.4
                     EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
                     SWISSPROT HIT: P34109, EVALUE 2.50e-01
                     NT HIT: gi11419444, EVALUE 3.00e-96"

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| ||||||||||    |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | ||||||||  ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1


RESULT 967
CQ350706/c
LOCUS       CQ350706                387 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24800 from Patent WO0157275.
ACCESSION   CQ350706
VERSION     CQ350706.1  GI:41299777
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 24800 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004909.2
                     EXPRESSED IN BRAIN, SIGNAL = 1.5
                     EST_HUMAN HIT: AV708482.1, EVALUE 1.00e-111
                     SWISSPROT HIT: P34109, EVALUE 2.50e-01
                     NT HIT: gi11419444, EVALUE 3.00e-96"

  Query Match             3.8%;  Score 285.7;  DB 1;  Length 387;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches  325;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        160 TTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCC 219
              || | | ||| |||||||||||||||||| ||| || | ||| ||| |||||| ||||||
Db        386 TTTCACTCCGCCCACCACTGCTGTTTGCCGCCATCGTAGACCCGCCTCTGACTTCCATCC 327

Qy        220 CTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCTC 279
              |||  ||||  |||||||||| |||| ||||||  |||||| |:| |||||||||| |||
Db        326 CTCCAGATCTGGCAGGGTGTCTGCTGCGCTCCTATTCCAGCCAGGGGCCCATTGCCCCTC 267

Qy        280 CCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAA 339
              || || |||||||||||||||||||||| |||||| ||||||||||    |||| || || |
Db        266 CCGATGGGGCTAAAGGCTTGCCATTGTTCCTGCAAGGCTAAGTGCTCAGGTTCGTCCTCA 207

Qy        340 TTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAAT 399
              |  |||||||  ||| |||||||||||||| |||||||||| ||||||||||||||||||
Db        206 TCAAGCTGAATCCTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTCTAAC 147

Qy        400 AKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAC 459
              |:  |||||||||| ||||||||||||||||||||||||||||||||||||||||:|||
Db        146 AGGGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGCCAA- 88

Qy        460 GAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCT 519
              |||| |||||||||||||| || |:|||||||||| | ||||||||  ||| ||| ||:| |
Db         87 GAACCCCAGGTCAGAGAACACAAGGCTTGCCACCGTATTGGAAGCAGCCCGCTGCCATTT 28

Qy        520 TGGAAGTGGTTCACCACCATCTTGGGA 546
              |||||||||    ||||||||||||||
Db         27 TGGAAGTGGCCTGCCACCATCTTGGGA 1
```

```
RESULT 968
AY208132
LOCUS       AY208132                329 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens isolate 1 endogenous retrovirus HERV-W long terminal
            repeat, partial sequence.
ACCESSION   AY208132
VERSION     AY208132.1  GI:28932683
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
   PUBMED   12867655
REFERENCE   2  (bases 1 to 329)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="1"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>329

  Query Match             3.8%;  Score 285.4;  DB 1;  Length 329;
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  292;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||||||| ||:|||||||||:|||||||||:|||||||||| |
Db             1 AGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGAAC 60

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
             |||||||||||||||||||| ||||||||||:|||||||  |||||||||||||||||
Db            61 GCATCCACCTTTAAACACGGGACTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTC 120

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
             |||||||| |||||||||||| |||||||||||||||||||||||||:||||||:||
Db           121 ACTAAAATACTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 180

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
             |||||||||| |||||||||| |||||| |||||||||||  |  || ||||||||||| |
Db           181 AGAGCACAGCGGGAGGGACAAGGATCGGCATATAAACCCAGGCATTGGAGCCGGCAACTG 240

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
             |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db           241 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCACTGTATTTCACTCT 300

Qy       7564 ATTAAATCTTGCARCTGCR 7582
             |||||||||||||:||||:
Db           301 ATTAAATCTTGCAACTGCA 319


RESULT 969
BD268834
LOCUS       BD268834                341 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268834
VERSION     BD268834.1  GI:33078602
KEYWORDS    JP 2002537846-A/13.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 341)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 13 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/13
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                  Location/Qualifiers
            FT   source          1. .341
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.8%;  Score 283.3;  DB 1;  Length 341;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  314;  Conservative    8;  Mismatches   10;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||:|||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||||||||||||||||||| ||||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTCGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  |||||||          |||| |||||||||||||||||||||||||||||||
Db           121 CACACCCAACCAATCAGGTAGTAAAGAGGGCTCACTAAAATGCTAATTAGGCAAAGACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||||||||:||||||:|||||||||||||||||||||||||||||
Db           181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||||||:|||||||||||| ||| |||  ||||||||||||||||||||||
Db           241 GGATATAAACCCAAGTCTTCGAGCCGGCAATGGCTACCTTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||||          |||||||||||||||||||||:||||
Db           301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 341


RESULT 970
AX035742
LOCUS       AX035742                341 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 13 from Patent DE19910650.
ACCESSION   AX035742
VERSION     AX035742.1  GI:11191325
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 13 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Human endogenous retrovirus"
```

```
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:11827"

  Query Match             3.8%;  Score 283.3;  DB 1;  Length 341;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  314;  Conservative    8;  Mismatches   10;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||:|||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||||||||||||||||||| |||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTCGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  |||||||         |||| ||||||||||||||||||||||||||||||
Db           121 CACACCCAACCAATCAGGTAGTAAAGAGGGCTCACTAAAATGCTAATTAGGCAAAGACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
Db           181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||||||||:|||||||||||||| ||| |||  ||||||||||||||||||
Db           241 GGATATAAACCCAAGTCTTCGAGCCGGCAATGGCTACCTTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||||||||||||          ||||||||||||||||||||:||||
Db           301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 341


RESULT 971
AB079270
LOCUS       AB079270                375 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pan troglodytes DNA, HERV-W LTR element, isolate:CH-2.
ACCESSION   AB079270
VERSION     AB079270.1  GI:28201101
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 375)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="CH-2"
                     /db_xref="taxon:9598"
                     /note="HERV-W LTR element"
     LTR             <1. .>375
                     /note="HERV-W LTR like element"

  Query Match             3.8%;  Score 281.4;  DB 1;  Length 375;
  Best Local Similarity   88.0%;  Pred. No. 0;
  Matches  329;  Conservative    5;  Mismatches   36;  Indels    4;  Gaps    3;

Qy        150 AGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTG 209
              |||||||||||| |||| || ||||||||||||||||||| |   |||| || ||| ||||
Db          4 AGCTGAGCTTTGGCTCGCCATCCACCACTGCTGTTTGCCGCNGTCGCACACNTGCTGCTG 63

Qy        210 ACTCCCATCCCTCTGGATCCTGCAGG--GTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||||||||||| || ||| |||||  |||||||||||||||||||||||||||:| ||
Db           64 ACTCCCATCCCTCCGGGTCCAGCAGGGTGTGTCCGCTGTGCTCCTGATCCAGCGAGGTGC 123

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||  ||||| |||||||||||| |||||| |||||||||||||||||| |
Db           124 CCATTGCCGCTCCTGATTGGACTAAAGGCTTGTCATTGTTCCTGCACGGCTAAGTGCCCG 183

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || |||||  |||||||||||| |||||||||||| ||||||||||  ||||||
Db           184 GGTTCGTCCTAATCCAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCTTGACCC 243

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              |||||||||||||:| |||||||||| ||| |||||||||||||||||||||||||||||
Db           244 ACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGAATC 303

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||||| ||||:|||||||||||||| ||||||||||
Db           304 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATC-TGGAAGCGGC 361

Qy        508 CTGCTACCRTCTTG 521
              |||| | |:|||||
Db           362 CTGCCAACATCTTG 375


RESULT 972
AB055869
LOCUS       AB055869                374 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:A431-1.
ACCESSION   AB055869
VERSION     AB055869.1  GI:15207778
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 374)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="A431-1"
                     /db_xref="taxon:9606"
                     /cell_line="A431"
     LTR             1. .374
                     /note="HERV-W family"

  Query Match             3.8%;  Score 281;  DB 1;  Length 374;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  323;  Conservative    5;  Mismatches   43;  Indels    2;  Gaps    2;

Qy        149 GAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG-TTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||  ||||| ||||||||||||| |||||||||  | ||  ||  ||||
Db          2 GAGCTGAGCTTTCCCTCACTGTCCACCACTGCTGTTTTGCCACCGTCACAGGCCCACCGC 61

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||| |||| | ||||||||  ||||| ||||| |||||||||||||||||||:||||
Db         62 TGACTTCCATTCTTCTGGATCTAGCAGGCTGTCCACTGTGCTCCTGATCCAGCGAGGCGC 121

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              ||||||||||||| |||||||||||||||||||||||| |||||| ||||| |||||||
```

```
Db        122 CCATTGCCGCTCCCGATTGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTACGTGCCTG 181

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||  | |||||  ||| |||||||| ||||||||||||| ||||||||||  ||||||
Db        182 GGTTCATCCTAATCAAGCCGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCATGACCC 241

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||| |||||||||:|||||||||||| ||| ||||||||||||||||||||||||||||
Db        242 ACGACTTCTAATAGAACTATAACACTCACCTCATGGCCCAAGATTCCATTCCTTGGAATC 301

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              | ||||||:||| |||| |||||||||||||| ||||:||||||||||||||||||| ||
Db        302 CATGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGTGGT 360

Qy        508 CTGCTACCRTCTT 520
                  | |||:||||
Db        361 TCACCACCATCTT 373


RESULT 973
AY208133
LOCUS       AY208133                328 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens isolate 2 endogenous retrovirus HERV-W long terminal
            repeat, partial sequence.
ACCESSION   AY208133
VERSION     AY208133.1  GI:28932684
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
   PUBMED   12867655
REFERENCE   2  (bases 1 to 328)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .328
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="2"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>328

  Query Match             3.7%;  Score 280.6;  DB 1;  Length 328;
  Best Local Similarity   90.6%;  Pred. No. 0;
  Matches  289;  Conservative    8;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
              |||||||||||||||||||||| ||: ||||||||:|||||||||:|| ||||||||||
Db          1 AGGACTAGCTGGATTTCCTAGGCCGACGAAGAATCCCTAAGCCTAGCTGAGAAGGTGACC 60

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
              |||||||||||||||| || |||||||||||:|||||||  ||||||||||||||||||
Db         61 GCATCCACCTTTAAACAGGGAGCTTGCAACTTAGTTCACACCCAACCAATCAGAGAGCTC 120

Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
              ||||||||| |||||| ||||| |||||||||||||||||||||||:||||||:||
Db        121 ACTAAAATGGTAATTAAGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG 180

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGG 7503
              |||||| ||| ||||||| || ||||||||||||||||||  |  ||||||| ||||||||
Db        181 AGAGCAGAGCTGGAGGGATAAGGATCGGGATATAAACCCAGGCATTCGAGCAGGCAACGG 240

Qy       7504 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
              ||||||||||||||||||||||||||||||||||||||| ||||||||||  ||||||||||||
Db        241 CAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGGGCTCTGTTTTCACTCTATTTCACTCT 300

Qy       7564 ATTAAATCTTGCARCTGCR 7582
              ||||||||||||||:||||:
Db        301 ATTAAATCTTGCAACTGCA 319


RESULT 974
AB055875
LOCUS       AB055875                374 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:RT4-2.
ACCESSION   AB055875
VERSION     AB055875.1  GI:15207784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 374)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="RT4-2"
                     /db_xref="taxon:9606"
                     /cell_line="RT4"
     LTR             1. .374
                     /note="HERV-W family"

  Query Match             3.7%;  Score 280.4;  DB 1;  Length 374;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  323;  Conservative    5;  Mismatches   44;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG-TTTGCCACCACCGCANACCTGCCG 206
              | |||||||||| ||||||| |||||||||||| |||||| |   | || ||| |||
Db          1 CCAGCTGAGCTTTCGCTCACCATCCACCACTGCTGTTTTGCCGCTGTCACAGACCCGCCA 60

Qy        207 CTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCG 266
              |||||| |||||||||||||||| |||||||| |||||||||||||||||||||| |:| |
Db         61 CTGACTTCCATCCCTCTGGATCCAGCAGGGTGCCCGCTGTGCTCCTGATCCAGCAAGGTG 120

Qy        267 CCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCT 326
              |||||||| ||||| |||||| |||||||||||||||| |||||| ||||||||||||
Db        121 CCCATTGCCACTCCCGATTGGGTTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCT 180

Qy        327 GGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACC 386
              |||||  | |||||  |||||||||||| |||||||||||| ||||||||||  |||||
Db        181 GGGTTCATCCTAATCCAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCATGACC 240

Qy        387 CACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 446
              |||  |||||||||:| |||||||||||||  |||||||||||||||||||||||||||||
Db        241 CACAACTTCTAATAGAGCTATAACACTTACTGCATGGCCCAAGATTCCATTCCTTGGAAT 300

Qy        447 CCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGG 506
              ||||||||:||| |||| |||||||||||||| ||||: ||||||||||||||||||| ||
Db        301 CCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGTCTTGCCACCATCTTGGAAGTGG 359

Qy        507 CCTGCTACCRTCTT 520
                    | |||:||||
Db        360 TTCACCACCATCTT 373
```

```
RESULT 975
AB055879
LOCUS       AB055879                374 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:UO31-2.
ACCESSION   AB055879
VERSION     AB055879.1  GI:15207788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Identification and phylogeny of new human endogenous retrovirus
            HERV-W family in cancer cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 374)
  AUTHORS   Lee,J. and Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmlee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="UO31-2"
                     /db_xref="taxon:9606"
                     /cell_line="UO-31"
     LTR             1. .374
                     /note="HERV-W family"

  Query Match             3.7%;  Score 280.4;  DB 1;  Length 374;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  323;  Conservative    5;  Mismatches   44;  Indels    2;  Gaps    2;

Qy        149 GAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG-TTTGCCACCACCGCANACCTGCCGC 207
              ||||||||||||  ||||| |||||||||||||| |||||||||  | ||  ||  ||||
Db          2 GAGCTGAGCTTTCCCTCACTGTCCACCACTGCTGTTTTGCCACCGTCACAGGCCCACCGC 61

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||  |||| | ||||||||  ||||| ||||| |||||||||||||||||||||:||||
Db         62 TGACTTCCATTCTTCTGGATCTAGCAGGCTGTCCACTGTGCTCCTGATCCAGCGAGGCGC 121

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||||||||| |||||||||||||||||||||||| |||||| ||||| |||||||
Db        122 CCATTGCCGCTCCCGATTGGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTACGTGCCTG 181

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              ||||  | |||||  ||| |||||||| |||||||||||| || |||||||  ||||||
Db        182 GGTTCATCCTAATCAAGCCGAACACTAGTCACTGGGTTCCACGGCTCTCTTCCATGACCC 241

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              ||| |||||||||:|||||||||||| ||| |||||||||||||||||||||||||||||
Db        242 ACGACTTCTAATAGAACTATAACACTCACCTCATGGCCCAAGATTCCATTCCTTGGAATC 301

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              | |||||:||| |||| |||||||||||| |||:||||||||||||||||||||| ||
Db        302 CATGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGTGGT 360

Qy        508 CTGCTACCRTCTTG 521
                | |||:|||||
Db        361 TCACCACCATCTTG 374


RESULT 976
AB079284
LOCUS       AB079284                371 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pongo pygmaeus DNA, HERV-W LTR element, isolate:OR-2.
ACCESSION   AB079284
VERSION     AB079284.1  GI:28201115
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 371)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .371
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="OR-2"
                     /db_xref="taxon:9600"
                     /note="HERV-W LTR element"
     LTR             <1. .>371
                     /note="HERV-W LTR like element"

  Query Match             3.7%;  Score 277.8;  DB 1;  Length 371;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  321;  Conservative    5;  Mismatches   45;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||||||| || ||||||||||||||||||||||   | || ||| |||||
Db          1 CGAGCTGAGCTTTTGCTCGCCATCCACCACTGCTGTTTGCCACTGTCTCAGACCCGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||||  ||||||||  ||||| |||||||||||||||||||||||||||||: | |
Db         61 TGACTCTCATCCCTCCAGATCCGGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGACTC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||| ||||| ||   ||||||||||||||||||| |||||| ||||||||||||
Db        121 CCATTGCCACTCCCGATCTTGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || ||||| |||||||||||||| |||||||||||  |||||||||  ||||||
Db        181 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCATGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              || ||||||||||:| |||||| ||| | | |||||||||||||||||||||||||||||
Db        241 ACAGCTTCTAATAGAGCTATAA-ACTCATCGCATGGCCCAAGATTCCATTCCTTGGAATC 299

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||| | ||:|||||||||||||||||||| ||
Db        300 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACATGAGGCTTGCCACCATCTTGGAAGTGGT 358

Qy        508 CTGCTACCRTCTT 520
                | |||:||||
Db        359 TCACCACCATCTT 371


RESULT 977
AF531174
LOCUS       AF531174                346 bp    DNA     linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens endogenous retrovirus W long terminal repeat, partial
            sequence.
ACCESSION   AF531174
VERSION     AF531174.1  GI:22297302
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
   PUBMED   12867655
REFERENCE   2  (bases 1 to 346)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: retrovirus W"
     LTR             <1. .>346

  Query Match             3.7%;  Score 276.5;  DB 1;  Length 346;
  Best Local Similarity   89.1%;  Pred. No. 0;
  Matches  301;  Conservative    6;  Mismatches   26;  Indels    5;  Gaps    1;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||| ||:|||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||  |||| ||||||||||||||||||||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACTGCATCTACCTTTAAACACGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC-----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGG 7414
              |||||| ||||||||     ||||||||||||||||  |||||||||||| ||||||||
Db           121 CACACCCGACCAATCAGTAAAGAGAGCTCACTAAAAATCTAATTAGGCAAAAACAGGAGG 180

Qy       7415 TAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGAT 7474
              ||||||||||||||||||:|| |||:||||||||||  |||||||| ||||| ||||||
Db           181 TAAAGAAATAGCCAATCATCTAATGCCTGAGAGCACAGTGGGAGGGACGATGATTGGGAT 240

Qy       7475 ATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATG 7534
              ||||||||| |  |||||||||||||| ||||||| ||||||||||||||||||||||||
Db           241 ATAAACCCAGGCATTCGAGCCGGCAACAGCAACCCTCTTTGGGTCCCCTCCCTTTGTATG 300

Qy       7535 GGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCT 7572
              ||| |||||||||||  |||||||||||| ||  | ||
Db           301 GGATCTCTGTTTTCACTCTATTTCACTCTTTTGCAACT 338


RESULT 978
AB079288
LOCUS       AB079288                372 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pongo pygmaeus DNA, HERV-W LTR element, isolate:OR-6.
ACCESSION   AB079288
VERSION     AB079288.1  GI:28201119
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 372)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="OR-6"
                     /db_xref="taxon:9600"
                     /note="HERV-W LTR element"
     LTR             <1. .>372
                     /note="HERV-W LTR like element"

  Query Match             3.7%;  Score 274.6;  DB 1;  Length 372;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  319;  Conservative    5;  Mismatches   47;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGC 207
              |||||||||||||| ||| | |||||||||||||||||||||||   | || ||| |||||
Db          1 CGAGCTGAGCTTTCGCTTGCTGTCCACCACTGCTGTTTGCCACTGTCTCAGACCCGCCGC 60

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              |||||| |||||||| ||||| ||||||||||||||||||||||||||||| ||: | |
Db         61 TGACTCTCATCCCTCCAGATCCGGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGACTC 120

Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              |||||||| ||||| ||   |||||||||||||||||| |||||| ||||||||||||
Db        121 CCATTGCCACTCCCGATCTTGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCTG 180

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || ||||| |||||||||||| |||||||||||  |||||||||  ||||||
Db        181 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCATGACCC 240

Qy        388 ACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATC 447
              || |||||||||:| |||||| ||| | | ||||||||||||||||||||||||||||||
Db        241 ACAGCTTCTAATAGAGCTATAA-ACTCATCGCATGGCCCAAGATTCCATTCCTTGGAATC 299

Qy        448 CGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGC 507
              ||||||:||| |||| |||||||||||| | ||:||||||||||||||||||| ||
Db        300 TGTGAGGCCAA-GAACCCCAGGTCAGAGAACATGAGGCTTGCCACCATCTTGGAAGTGGT 358

Qy        508 CTGCTACCRTCTT 520
                | |||:||||
Db        359 TCACCACCATCTT 371


RESULT 979
AB071971
LOCUS       AB071971                373 bp    DNA     linear   PRI 23-MAR-2002
DEFINITION  Homo sapiens DNA, HERV-W LTR, isolate:W-Sch2.
ACCESSION   AB071971
VERSION     AB071971.1  GI:19698206
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,H. and Huh,J.
  TITLE     Isolation of HERV-W LTR Elements in Schizophrenia
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2001) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="W-Sch2"
                     /db_xref="taxon:9606"
     LTR             1. .373
```

```
                    /note="HERV-W LTR"

  Query Match             3.7%;  Score 273.8;  DB 1;  Length 373;
  Best Local Similarity   87.5%;  Pred. No. 0;
  Matches  314;  Conservative    4;  Mismatches   39;  Indels    2;  Gaps    2;

Qy        148 CGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG-TTTGCCACCACCGCANACCTGCCG 206
              | |||||||||||| ||||||| |||||||||||||| |||||| |    | || ||| |||
Db          1 CCAGCTGAGCTTTCGCTCACCATCCACCACTGCTGTTTTGCCGCTGTCACAGACCCGCCA 60

Qy        207 CTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCG 266
              |||||| |||||||||||||||| |||||||| |||||||||||||||||||||| |:| |
Db         61 CTGACTTCCATCCCTCTGGATCCAGCAGGGTGCCCGCTGTGCTCCTGATCCAGCAAGGTG 120

Qy        267 CCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCT 326
              ||||||||| ||||| |||||| || |||||||||||||| |||||| ||||||||||||
Db        121 CCCATTGCCACTCCCGATTGGGTTAGAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCT 180

Qy        327 GGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACC 386
              |||||  | |||||  |||||||||||| |||||||||||| ||||||||||  |||||
Db        181 GGGTTCATCCTAATCCAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCATGACC 240

Qy        387 CACGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAAT 446
              |||  |||||||||:| |||||||||||||  |||||||||||||||||||||||||||||
Db        241 CACAACTTCTAATAGAGCTATAACACTTACTGCATGGCCCAAGATTCCATTCCTTGGAAT 300

Qy        447 CCGTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCG 505
              ||||||||:||| |||| |||||||||||||| ||||: ||||||||||||||||||||
Db        301 CCGTGAGGCCAA-GAACCCCAGGTCAGAGAACACGAGTCTTGCCACCATCTTGGAAGCG 358


RESULT 980
AB079287
LOCUS       AB079287                373 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Pongo pygmaeus DNA, HERV-W LTR element, isolate:OR-5.
ACCESSION   AB079287
VERSION     AB079287.1  GI:28201118
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Huh,J. and Kim,H.
  TITLE     Isolation and Phylogeny of HERV-W LTR Family in Hominoid Primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 373)
  AUTHORS   Kim,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="OR-5"
                     /db_xref="taxon:9600"
                     /note="HERV-W LTR element"
     LTR             <1. .>373
                     /note="HERV-W LTR like element"

  Query Match             3.6%;  Score 271.7;  DB 1;  Length 373;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches  311;  Conservative    5;  Mismatches   56;  Indels    1;  Gaps    1;

Qy        149 GAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCT 208
              |||||||||||||| || || |||||||||||||||||| ||| ||| | ||||||||
Db          2 GAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGCCGCCATTGCAGAGCTGCCGCT 61

Qy        209 GACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCC 268
              |||| |||||||||  ||||  |||| |||||| |||||||||||||||||:|  ||
Db         62 GACTTCCATCCCTCCAGATCTGGCAGAGTGTCCACTGTGCTCCTGATCCAGTGAGGGTCC 121

Qy        269 CATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGG 328
              ||||||| |||| ||  |||||||||||||||||||| ||||  ||||||||||| ||
Db        122 CATTGCCACTCCTGATCAGGCTAAAGGCTTGCCATTGTTCCTGTGTGGCTAAGTGCCCGG 181

Qy        329 GTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCA 388
              |||  | |||||  |||||||||||| |||||| ||||||  |||||||||  |||||||
Db        182 GTTCATCCTAATCAAGCTGAACACTAGTCACTGCGTTCCACAGTTCTCTTCCATGACCCA 241

Qy        389 CGGCTTCTAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCC 448
              | ||||||||||:|  ||||||||| ||  |||||||||||||||||||||||||||||
Db        242 CAGCTTCTAATAGAGGTATAACACTCACTGCATGGCCCAAGATTCCATTCCTTGGAATCC 301

Qy        449 GTGAGGSCAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCC 508
              |||||:||| |||||||| |||||||| || |:|||||||||| |||||||| ||||
Db        302 ATGAGGCCAA-GAACTCCATGTCAGAGAACACCAGGCTTGCCACCTTCTTGGAAGTGGCC 360

Qy        509 TGCTACCRTCTTG 521
              |  ||:| |||
Db        361 CACCGCCATTTTG 373


RESULT 981
BD268847
LOCUS       BD268847                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268847
VERSION     BD268847.1  GI:33078615
KEYWORDS    JP 2002537846-A/26.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 26 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/26
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                     Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.6%;  Score 270.3;  DB 1;  Length 343;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  307;  Conservative    7;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||| |||||||||| |||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAAACAGGACTAGCAGGATTTCCTAGGCCGATTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||| ||||||| |||||||||||||||||||||||||||||||||||| ||:|
Db         61 CTAAGCCTAGATGGGAAGTTGACCACATCCACCTTTAAACACGGGGCTTGCAACTCAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
```

```
                ||||||  ||||  |||          |||||||  |  ||||||||||||||||||||  ||||
Db        121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
                ||||||||||||||||||||||:||||||:|||||||||||  ||||||||||||||
Db        181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                |||||||||||||  |  ||||||||||||||  | ||||  ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                |||||||||||||||          |||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 982
AX035755
LOCUS       AX035755                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 26 from Patent DE19910650.
ACCESSION   AX035755
VERSION     AX035755.1  GI:11191338
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 26 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.6%;  Score 270.3;  DB 1;  Length 343;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  307;  Conservative    7;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
                |||||||||||||||||||| |||||||||| ||||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAAACAGGACTAGCAGGATTTCCTAGGCCGATTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
                :|||||||| ||||||| ||||||||||||||||||||||||||||||||||| ||:|
Db         61 CTAAGCCTAGATGGGAAGTTGACCACATCCACCTTTAAACACGGGGCTTGCAACTCAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
                ||||||  ||||  |||          |||||||  |  ||||||||||||||||||||  ||||
Db        121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
                |||||||||||||||||||||:||||||:|||||||||||  ||||||||||||||
Db        181 GAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                |||||||||||||  |  ||||||||||||||  | ||||  ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                |||||||||||||||          |||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 983
BD268845
LOCUS       BD268845                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268845
VERSION     BD268845.1  GI:33078613
KEYWORDS    JP 2002537846-A/24.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 24 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/24
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                  Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.6%;  Score 269.9;  DB 1;  Length 343;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  306;  Conservative    8;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
                |||||||||||||||||||||||||||||||| ||||||||||||||  |:|||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCAACTAAGAATCT 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
                :|||||||||:||||||||||||||||||||||||||||||| ||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
                ||||||  ||||  |||          |||||||  |  ||||||||||||||||||||  ||||
Db        121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGTCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 7470
                |||||||||||||||  |||||||:||||||:|||||||||||||  ||||||||||||
Db        181 GAGGTAAAGAAATAGTCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                |||||||||||||  |  ||||||||||||||  | ||||  ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                |||||||||||||||          |||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 984
AX035753
LOCUS       AX035753                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 24 from Patent DE19910650.
ACCESSION   AX035753
VERSION     AX035753.1  GI:11191336
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 24 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
```

```
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.6%;  Score 269.9;  DB 1;  Length 343;
  Best Local Similarity   86.9%;  Pred. No. 0;
  Matches  306;  Conservative    8;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||| ||||||||||||||  |:||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCAACTAAGAATCT 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||| ||||||||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| |||| |||         ||||||| | ||||||||||||||||||||| ||||
Db        121 CACACCCGACCCATCAGGTAAGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||| |||||||:||||||:|||||||||||| ||||||||||||||
Db        181 GAGGTAAAGAAATAGTCAATCATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||| |  |||||||||||||| | |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 985
CQ052449/c
LOCUS       CQ052449                379 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 3269 from Patent WO0157270.
ACCESSION   CQ052449
VERSION     CQ052449.1  GI:41026955
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 3269 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN HBL100, SIGNAL = 1.4"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             |||||||||||  ||   ||| |      ||  |||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| ||||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||

Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| |||||||||||  ||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |  | |||||||||||||||||||||||| ||||  ||||||||
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
             |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db         85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
             | |  |  |||||| || | |
Db         25 GGACTAGCCCCTGCAATACTCC 4


RESULT 986
CQ067518/c
LOCUS       CQ067518                379 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 3318 from Patent WO0157278.
ACCESSION   CQ067518
VERSION     CQ067518.1  GI:41037387
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 3318 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN HELA, SIGNAL = 1.6"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             |||||||||||  ||   ||| |      ||  |||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| ||||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||

Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| |||||||||||  ||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |  | |||||||||||||||||||||||| ||||  ||||||||
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
             |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db         85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26
```

```
Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
Db         25 GGACTAGCCCCTGCAATACTCC 4


RESULT 987
CQ094566/c
LOCUS       CQ094566                379 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3425 from Patent WO0157272.
ACCESSION   CQ094566
VERSION     CQ094566.1  GI:41063592
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 3425 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN PLACENTA, SIGNAL = 3.3"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
Db         85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
Db         25 GGACTAGCCCCTGCAATACTCC 4


RESULT 988
CQ133333/c
LOCUS       CQ133333                379 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3355 from Patent WO0157276.
ACCESSION   CQ133333
VERSION     CQ133333.1  GI:41090689
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 3355 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.3"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
Db         85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
Db         25 GGACTAGCCCCTGCAATACTCC 4


RESULT 989
CQ201049/c
LOCUS       CQ201049                379 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3326 from Patent WO0157271.
ACCESSION   CQ201049
VERSION     CQ201049.1  GI:41186713
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 3326 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN BT474, SIGNAL = 1.3"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| |||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |   | ||||||||||||||||||||| ||||   |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 990
CQ216554/c
LOCUS       CQ216554                379 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 3393 from Patent WO0157273.
ACCESSION   CQ216554
VERSION     CQ216554.1  GI:41197878
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 3393 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .379
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN ADULT LIVER, SIGNAL = 2"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
             |||||||||  |||||||||||||||||||| ||| |||||||||| ||||||||||||
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544
             |||||||||||||| |   | ||||||||||||||||||||| ||||   |||||||| |
Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604
              |||||||||| |||||||| ||||| ||||| |||||||||||||| || |||||||||
Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626
              | |  |  ||||||| || | |
Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 991
CQ255127/c
LOCUS       CQ255127                379 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3388 from Patent WO0157277.
ACCESSION   CQ255127
VERSION     CQ255127.1  GI:41227605
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 3388 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .379
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN FETAL LIVER, SIGNAL = 1.2"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304
             ||||||||||||  ||   ||| |      ||  |||||||||||||||||||||| | |
Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364
             |||||||||||| |||| |||||||||||||||||||| |||||||| ||||||||||
Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424
             |||||||||||||||||||||||| ||||||| | | ||||| ||||||||||||||||||
Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484
```

```
Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544

Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604

Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626

Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 992
CQ292212/c
LOCUS       CQ292212                379 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3317 from Patent WO0186003.
ACCESSION   CQ292212
VERSION     CQ292212.1  GI:41252789
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 3317 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN LUNG, SIGNAL = 1.2"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304

Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364

Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424

Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484

Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544

Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604

Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626

Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 993
CQ329245/c
LOCUS       CQ329245                379 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 3339 from Patent WO0157275.
ACCESSION   CQ329245
VERSION     CQ329245.1  GI:41277812
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 3339 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN BRAIN, SIGNAL = 1.2"

  Query Match             3.6%;  Score 268.4;  DB 1;  Length 379;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  329;  Conservative    0;  Mismatches   47;  Indels    6;  Gaps    3;

Qy       2245 AGCCACTAAGTTATGAACTGGGGAGCTTTATTCTTTTCACATGCTTTTCTAATTATGCTT 2304

Db        379 AGCCACTAAGTTGGGACTGGGGAA----CTTTACTTTCACATGCTTTTCTAATTATACCT 324

Qy       2305 GAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACAC 2364

Db        323 GAAAGCCCCACTCCCTTATTAGGGAGAGACATTCTAGCAGAAGCAGGGGTCATTATACAC 264

Qy       2365 CTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCC 2424

Db        263 CTGAACATAGGAGAAGGAACACCTGTTTGTT-TCCACCTGCATGAGGAAGGAATTAATCC 205

Qy       2425 TGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAG 2484

Db        204 TGAAGTCTGATCAACAGAAGGACAATATGGATGAG-CAAAGAATGCCTGTCCTGTTCAAG 146

Qy       2485 TTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGG 2544

Db        145 TTAAACTAAAGGATGCTGCCTCCTTTCCCTACCAAAGGCAGTATCCCCATAGACCCAAAG 86

Qy       2545 CCCAACAAGGATTCCAAAAGATTGTTAAGGACTTAAAAGCCCAAGGCTTAGTAAAACCAT 2604

Db          85 TCCAACAAGGACTCCAAAAGGTTGTTGAGGACCTAAAAGCCCAAGGCCTAATAAAACCAT 26

Qy       2605 GCATAACTCCCTGCAGTAATTC 2626

Db          25 GGACTAGCCCCTGCAATACTCC 4


RESULT 994
BD268841
LOCUS       BD268841                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268841
VERSION     BD268841.1  GI:33078609
KEYWORDS    JP 2002537846-A/20.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
```

```
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 20 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/20
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 265.1;  DB 1;  Length 343;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  303;  Conservative    8;  Mismatches   22;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||| ||| |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         ||||||||  |||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||| ||||||||||||||:||| ||:||||||||||||||||||||||||||||||
Db        181 GAGGTAGAGAAATAGCCAATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCC 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||||||  |||||||| |||| ||| |||| ||||| |||||||||||||||
Db        241 GGATATAAACCCAAGCATTCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||| |||||          |||||||||||||||||:||||:
Db        301 TATGGGAGCTCTATTTTC----------ACTCTATTAAATCTTGCAACTGCA 342


RESULT 995
AX035749
LOCUS       AX035749                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 20 from Patent DE19910650.
ACCESSION   AX035749
VERSION     AX035749.1  GI:11191332
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 20 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 265.1;  DB 1;  Length 343;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  303;  Conservative    8;  Mismatches   22;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||| ||| |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         ||||||||  |||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||| ||||||||||||||:||| ||:||||||||||||||||||||||||||||||
Db        181 GAGGTAGAGAAATAGCCAATCATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCC 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||||||  |||||||| |||| ||| |||| ||||| |||||||||||||||
Db        241 GGATATAAACCCAAGCATTCGAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||| |||||          |||||||||||||||||:||||:
Db        301 TATGGGAGCTCTATTTTC----------ACTCTATTAAATCTTGCAACTGCA 342


RESULT 996
AY208136
LOCUS       AY208136                328 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens isolate 5 endogenous retrovirus HERV-W long terminal
            repeat, partial sequence.
ACCESSION   AY208136
VERSION     AY208136.1  GI:28932687
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Analysis of transcriptional regulatory sequences in the human
            endogenous retrovirus W long terminal repeat
  JOURNAL   J. Gen. Virol. 84 (Pt 8), 2229-2235 (2003)
  PUBMED    12867655
REFERENCE   2  (bases 1 to 328)
  AUTHORS   Lee,W.J., Kwun,H.J. and Jang,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2002) Microbiology, Pusan National University,
            #30 Changjeon-Dong, Kumjeong-Ku, Pusan 609-735, Korea
FEATURES             Location/Qualifiers
     source          1. .328
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="5"
                     /db_xref="taxon:9606"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>328

  Query Match             3.5%;  Score 263.4;  DB 1;  Length 328;
  Best Local Similarity   90.9%;  Pred. No. 0;
  Matches  291;  Conservative    8;  Mismatches   19;  Indels    2;  Gaps    2;

Qy       7264 AGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323
             |||||||||||||||||||||| ||:|||||||||:|||||||||:|||||||||||
Db          1 AGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 60

Qy       7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTC 7383
             ||||||||||||||||||||||||||||||:|||||||  ||||||||||||||||
Db         61 GCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGAGAGCTC 120
```

```
Qy       7384 ACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTG 7443
             |||||||||||||||||||||||| |||||||||||||||||||||||||||| :||||||:|
Db        121 ACTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCAACTATTGCCTT 180

Qy       7444 AGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAAC-G 7502
             |||||||||  |||||||||| ||||||||||||||||||| |  |||||||||||||| |
Db        181 AGAGCACAGAGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCCGGCAACAG 240

Qy       7503 GCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTC 7562
             ||||||| |||||||||||||||||||| | ||||||||||||||  ||||||||||||
Db        241 GCAACCCTCTTTGGGTCCCCTCCCTTTGTGT-GGAGCTCTGTTTTCACTCTATTTCACTC 299

Qy       7563 TATTAAATCTTGCARCTGCR 7582
             |||||||| ||||:||||:
Db        300 TATTAAATCATGCAACTGCA 319


RESULT 997
BD268835
LOCUS       BD268835                341 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268835
VERSION     BD268835.1  GI:33078603
KEYWORDS    JP 2002537846-A/14.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 341)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 14 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/14
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .341
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 262.9;  DB 1;  Length 341;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  302;  Conservative    7;  Mismatches   23;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||:||||||| |
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATTC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||| ||||||||| |||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              || | ||||||||||         ||||||||||||||||||||||||||||| || ||||
Db        121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||| |||||| ||||||||||:| ||||:||| |||||||||||||||||||||||||
Db        181 GAGGCAAAGAAGTAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  |||||||| || || || |||| |||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||          |||||||||||||||||||||:||||
Db        301 TATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGC 341


RESULT 998
AX035743
LOCUS       AX035743                341 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 14 from Patent DE19910650.
ACCESSION   AX035743
VERSION     AX035743.1  GI:11191326
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 14 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 262.9;  DB 1;  Length 341;
  Best Local Similarity   86.0%;  Pred. No. 0;
  Matches  302;  Conservative    7;  Mismatches   23;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATTC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||:||||||| |
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATTC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||| ||||||||| |||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              || | ||||||||||         ||||||||||||||||||||||||||||| || ||||
Db        121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||| |||||| ||||||||||:| ||||:||| |||||||||||||||||||||||||
Db        181 GAGGCAAAGAAGTAGCCAATCATCTGTTGCCTGACAGCACAGCAGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  |||||||| || || || |||| |||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCGAGCCAGCTACAGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              |||||||||||||          |||||||||||||||||||||:||||
Db        301 TATGGGAGCTCTGT----------CTTCACTCTATTAAATCTTGCAACTGC 341


RESULT 999
CQ055957
LOCUS       CQ055957                310 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 6777 from Patent WO0157270.
ACCESSION   CQ055957
VERSION     CQ055957.1  GI:41030463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
```

```
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 6777 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HBL100, SIGNAL = 1.8
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              |||||||||||||||||||||||| | |||||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 1000
CQ075199
LOCUS       CQ075199                310 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 10999 from Patent WO0157278.
ACCESSION   CQ075199
VERSION     CQ075199.1  GI:41045068
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 10999 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HELA, SIGNAL = 1.7
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
               | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              |||||||||||||||||||||||| | |||||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              ||||||||||||
Db        300 AGCAGGACTGA 310


RESULT 1001
CQ106152
LOCUS       CQ106152                310 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 15011 from Patent WO0157272.
ACCESSION   CQ106152
VERSION     CQ106152.1  GI:41075205
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 15011 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN PLACENTA, SIGNAL = 10
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| ||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
             |||||||||||||||||||||||||||| || | ||||||||||||||||||||||||
```

```
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970

Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030

Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041

Db        300 AGCAGGACTGA 310


RESULT 1002
CQ144842
LOCUS       CQ144842                310 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14864 from Patent WO0157276.
ACCESSION   CQ144842
VERSION     CQ144842.1  GI:41102214
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 14864 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.5
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790

Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850

Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910

Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970

Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030

Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041

Db        300 AGCAGGACTGA 310


RESULT 1003
CQ180280
LOCUS       CQ180280                310 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 11676 from Patent WO0157274.
ACCESSION   CQ180280
VERSION     CQ180280.1  GI:41175019
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 11676 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN HEART, SIGNAL = 1.2
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790

Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850

Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910

Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970

Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030

Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041

Db        300 AGCAGGACTGA 310


RESULT 1004
CQ204678
LOCUS       CQ204678                310 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 6955 from Patent WO0157271.
ACCESSION   CQ204678
VERSION     CQ204678.1  GI:41190663
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
```

```
          analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 6955 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN BT474, SIGNAL = 0.99
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| |||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
              |||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              |||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              |||||||||||
Db        300 AGCAGGACTGA 310


RESULT 1005
CQ228028
LOCUS       CQ228028                310 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14867 from Patent WO0157273.
ACCESSION   CQ228028
VERSION     CQ228028.1  GI:41211246
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 14867 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| |||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
             ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db         60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy       1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
              |||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db        120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy       1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy       1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
              |||||||||||||||||||||| | |||||||||||||||||||||| ||||||||
Db        240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy       2031 AGCAGGACTGA 2041
              |||||||||||
Db        300 AGCAGGACTGA 310


RESULT 1006
CQ266186
LOCUS       CQ266186                310 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 14447 from Patent WO0157277.
ACCESSION   CQ266186
VERSION     CQ266186.1  GI:41238790
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 14447 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy       1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
             |||||||||||||||| |||||||||||||||||| |||||| |||||||||||||||||
Db          1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy       1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
```

```
                    ||||| || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db          60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy        1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
                ||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db          120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy        1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db          180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy        1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                ||||||||||||||||||||||| | |||||||||||||||||||| ||||||||
Db          240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy        2031 AGCAGGACTGA 2041
                ||||||||||||
Db          300 AGCAGGACTGA 310


RESULT 1007
CQ303217
LOCUS       CQ303217                310 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 14322 from Patent WO0186003.
ACCESSION   CQ303217
VERSION     CQ303217.1  GI:41263794
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 14322 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN LUNG, SIGNAL = 1.6
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy        1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
                ||||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db           1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy        1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
                |||||  || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db          60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy        1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
                ||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db          120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy        1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db          180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy        1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                ||||||||||||||||||||||| | |||||||||||||||||||| ||||||||
Db          240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy        2031 AGCAGGACTGA 2041
                ||||||||||||
Db          300 AGCAGGACTGA 310


RESULT 1008
CQ340460
LOCUS       CQ340460                310 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 14554 from Patent WO0157275.
ACCESSION   CQ340460
VERSION     CQ340460.1  GI:41289531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 14554 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC004897.1
                     EXPRESSED IN BRAIN, SIGNAL = 0.91
                     NT HIT: AL163280.2, EVALUE 1.00e-105
                     EST_HUMAN HIT: AA668227.1, EVALUE 2.00e-23
                     SWISSPROT HIT: P14077, EVALUE 1.00e-09"

  Query Match             3.5%;  Score 262.7;  DB 1;  Length 310;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  287;  Conservative    0;  Mismatches   23;  Indels    1;  Gaps    1;

Qy        1731 AGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGATTAAAAAAAAGGCCACCGCTTTAG 1790
                ||||||||||||||||| |||||||||||||||| |||||| ||||||||||||||||
Db           1 AGATCAGGAGGAGCAGGTGGAACAGGACAAACGGGA-TAAAAAGAAGGCCACCGCTTTAG 59

Qy        1791 TCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATT 1850
                |||||  || ||||||||| |||||||||| |||||||| ||||||| ||||||||||
Db          60 TCATGGCCTTCAGGCAAGCGGACTTTGGAGACTCTGGAACAGGGAAAGGCTGGGCAAATC 119

Qy        1851 GAATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTC 1910
                ||||||||||||||||||||||||||| || | |||||||||||||||||||||||||
Db          120 AAATGCCTAATAGGGCTTGCTTCCAGTGCAGTTTGCAAGGACACTTTAAAAAAGATTGTC 179

Qy        1911 CAAGTAGAAGTAAGCCGCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGG 1970
                | ||||| |||||||||| |||||||||||||||| |||||||||||||||||||||
Db          180 AGAATAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAAGGGAATCACTGGAAGG 239

Qy        1971 CCCACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGC 2030
                ||||||||||||||||||||||| | |||||||||||||||||||| ||||||||
Db          240 TCCACTGCCCCAGGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACCAGAAGATCCAGC 299

Qy        2031 AGCAGGACTGA 2041
                ||||||||||||
Db          300 AGCAGGACTGA 310


RESULT 1009
BD268839
LOCUS       BD268839                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268839
VERSION     BD268839.1  GI:33078607
KEYWORDS    JP 2002537846-A/18.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
```

```
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 18 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/18
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||| :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||| |||||| |||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         || |||||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||| ||||||||||:||| ||:|||||||| |||||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||| ||||          |||||||||||||||||||:||||:
Db        301 TATGGGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1010
BD268844
LOCUS       BD268844                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268844
VERSION     BD268844.1  GI:33078612
KEYWORDS    JP 2002537846-A/23.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 23 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/23
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||| :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||| |||||| |||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||| ||||||||         || |||||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||| ||||||||||:||| ||:|||||||| |||||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||| ||||          |||||||||||||||||||:||||:
Db        301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1011
AX035747
LOCUS       AX035747                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 18 from Patent DE19910650.
ACCESSION   AX035747
VERSION     AX035747.1  GI:11191330
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 18 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||| :|||||||||
```

```
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| ||||||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  ||||||||         || ||||||||||||||||||||||| ||| ||||
Db         121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||  ||||||||||:||| ||:|||||||||| ||||||||||||||||||
Db         181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db         241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||| ||||          |||||||||||||||||||||:||||:
Db         301 TATGGGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1012
AX035752
LOCUS       AX035752                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 23 from Patent DE19910650.
ACCESSION   AX035752
VERSION     AX035752.1  GI:11191335
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 23 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 260.3;  DB 1;  Length 343;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  300;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||||||||||||| ||||||||||||||| :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| ||||||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  ||||||||         || ||||||||||||||||||||||| ||| ||||
Db         121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||  ||||||||||:||| ||:|||||||||| ||||||||||||||||||
Db         181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db         241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||  ||||||||         |||||||||||||||||||||:||||:
Db         301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1013
BD268840
LOCUS       BD268840                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268840
VERSION     BD268840.1  GI:33078608
KEYWORDS    JP 2002537846-A/19.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 19 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/19
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 258.7;  DB 1;  Length 343;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  299;  Conservative    8;  Mismatches   26;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||||||||||||| |||||||||||||||| | :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCCGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| ||||||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||||  ||||||||         || ||||||||||||||||||||||| ||| ||||
Db         121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||  ||||||||||:||| ||:|||||||||| ||||||||||||||||||
Db         181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||| |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db         241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||  ||||||||         |||||||||||||||||||||:||||:
Db         301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1014
AX035748
LOCUS       AX035748                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 19 from Patent DE19910650.
ACCESSION   AX035748
VERSION     AX035748.1  GI:11191331
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
```

```
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 19 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.5%;  Score 258.7;  DB 1;  Length 343;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  299;  Conservative    8;  Mismatches   26;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| |||||||||||||| | :|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGTTGGATTTCCTAGGCCGGCTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:||||||  ||||||| |||||||||||||||||||||||| ||||:|
Db         61 CTAAGCCTAGCTGGGAAATTGACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              |||||| ||||||||         || ||||||||||||||||||||||| ||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||| |||||||||:||| ||:|||||||||| ||||||||||||||||||
Db        181 GAGGTAAAGAAGTAGCCAATCATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||||||  |  ||| |||| ||   ||| |||| ||||||||||||||||||||
Db        241 GGATATAAACCCAGGCATTCAAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTG 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||  ||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGAAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1015
BD268842
LOCUS       BD268842                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268842
VERSION     BD268842.1  GI:33078610
KEYWORDS    JP 2002537846-A/21.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 21 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/21
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                  Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| |||||| |||||||| |||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              || | |||||||||          ||||||||  ||||||||||||||||||||| ||||
Db        121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||||||  ||||:||| ||:||| ||||||    || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1016
BD268843
LOCUS       BD268843                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268843
VERSION     BD268843.1  GI:33078611
KEYWORDS    JP 2002537846-A/22.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 22 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/22
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                  Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
                 :|||||||||:||||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
                 |||||| ||||||||         ||||||||  |||||||||||||||||||||| ||||
Db         121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
                 |||||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db         181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                 |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db         241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                 ||||||||||||||         |||||||||||||||||||||:||||:
Db         301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1017
AX035750
LOCUS       AX035750                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 21 from Patent DE19910650.
ACCESSION   AX035750
VERSION     AX035750.1  GI:11191333
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 21 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
                 ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
                 :|||||||||:||||||||||||| |||||| |||||||| ||||||||||||||| :|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
                 || | ||||||||||         ||||||||  |||||||||||||||||||||| ||||
Db         121 CATATCTGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
                 |||||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db         181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                 |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db         241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                 ||||||||||||||         |||||||||||||||||||||:||||:
Db         301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1018
AX035751
LOCUS       AX035751                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 22 from Patent DE19910650.
ACCESSION   AX035751
VERSION     AX035751.1  GI:11191334
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 22 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 255.5;  DB 1;  Length 343;
  Best Local Similarity   84.4%;  Pred. No. 0;
  Matches  297;  Conservative    8;  Mismatches   28;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
                 ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
                 :|||||||||:||||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
                 |||||| ||||||||         ||||||||  |||||||||||||||||||||| ||||
Db         121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
                 |||||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db         181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
                 |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db         241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
                 ||||||||||||||         |||||||||||||||||||||:||||:
Db         301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1019
BD268837
LOCUS       BD268837                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268837
VERSION     BD268837.1  GI:33078605
KEYWORDS    JP 2002537846-A/16.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 16 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/16
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
```

```
           PC   A61P43/00,
           PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
           Retrovirus expression vector based on HERVLTR sequence FH   Key
                Location/Qualifiers
           FT   source          1. .343
           FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 253.9;  DB 1;  Length 343;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  296;  Conservative    8;  Mismatches   29;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||||| ||||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAAACATGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             ||||||  |||||||          ||||||||  ||||||||||||||||||||  ||||
Db        121 CACACCCAACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             ||||||||||||||||| ||||:||| ||:||| ||||||    || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||  ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1020
AX035745
LOCUS       AX035745                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 16 from Patent DE19910650.
ACCESSION   AX035745
VERSION     AX035745.1  GI:11191328
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 16 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 253.9;  DB 1;  Length 343;
  Best Local Similarity   84.1%;  Pred. No. 0;
  Matches  296;  Conservative    8;  Mismatches   29;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||||| ||||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAAACATGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             ||||||  |||||||          ||||||||  ||||||||||||||||||||  ||||
Db        121 CACACCCAACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             ||||||||||||||||| ||||:||| ||:||| ||||||    || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||  ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1021
BD268846
LOCUS       BD268846                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268846
VERSION     BD268846.1  GI:33078614
KEYWORDS    JP 2002537846-A/25.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 25 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/25
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 252.3;  DB 1;  Length 343;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  295;  Conservative    8;  Mismatches   30;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||| ||| |||||||||| ||| |||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             ||||||  |||||||          ||||||||  ||||||||||||||||||||  ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             ||||||||||||||||| ||||:||| ||:||| ||||||    || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240
```

```
Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1022
AX035754
LOCUS       AX035754                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 25 from Patent DE19910650.
ACCESSION   AX035754
VERSION     AX035754.1  GI:11191337
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 25 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.4%;  Score 252.3;  DB 1;  Length 343;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches  295;  Conservative    8;  Mismatches   30;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||  |:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||| ||| |||||||||| |||  |||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              |||||| ||||||||          ||||||| ||||||||||||||||||||| ||||
Db        121 CACACCCGACCAATCAGGTAGTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              ||||||||||||||||| ||||:||| ||:||| ||||||   || ||||||||||||
Db        181 GAGGTAAAGAAATAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1023
BD268838
LOCUS       BD268838                343 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268838
VERSION     BD268838.1  GI:33078606
KEYWORDS    JP 2002537846-A/17.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 17 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/17
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .343
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.3%;  Score 250.7;  DB 1;  Length 343;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  294;  Conservative    8;  Mismatches   31;  Indels   19;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTCCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              || |||  || ||||          |||||||||||||||||| || ||||||| || ||||
Db        121 CATACCCAACAAATCAGGTAGTAAAGAGAGCTCACTAAAATACTGATTAGGCGAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              ||||||| |||| ||||| ||||:||| ||:||| ||||||    || ||||||||||||
Db        181 GAGGTAAGGAAACAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
              |||||||||||||          |||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342


RESULT 1024
AX035746
LOCUS       AX035746                343 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 17 from Patent DE19910650.
ACCESSION   AX035746
VERSION     AX035746.1  GI:11191329
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 17 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"

  Query Match             3.3%;  Score 250.7;  DB 1;  Length 343;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches  294;  Conservative    8;  Mismatches   31;  Indels   19;  Gaps    2;
```

```
Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTCCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             || |||  || ||||         |||||||||||||||| || ||||||| || ||||
Db        121 CATACCCAACAAATCAGGTAGTAAAGAGAGCTCACTAAAATACTGATTAGGCGAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             ||||||| |||| ||||| |||||||||| ||:||| |||||| || ||||||||||||
Db        181 GAGGTAAGGAAACAGCCAGTCATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCA 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             |||||||||| || |  ||| |||| |||| ||| |||| ||||||||||||||||| ||
Db        241 GGATATAAACTCAGGCATTCAAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTT 300

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGCR 7582
             ||||||||||||||        ||||||||||||||||||||||:||||:
Db        301 TATGGGAGCTCTGT----------TTTCACTCTATTAAATCTTGCAACTGCA 342
```

RESULT 1025
CQ741068/c
LOCUS CQ741068 300 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 27002 from Patent WO02068579.
ACCESSION CQ741068
VERSION CQ741068.1 GI:42348218
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 27002 06-SEP-2002; PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .300
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 3.3%; Score 246; DB 1; Length 300;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 264; Conservative 1; Mismatches 34; Indels 0; Gaps 0;

```
Qy         90 CTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCG 149
              |||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||| |
Db        300 CTATTTCACTCTATTAAATCTTGTAACTGCACTCTTCTGGTCCGTGTTTGTTACGGCTTG 241

Qy        150 AGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTG 209
              |||||||||||| |||||||||||||| |||||||||||| ||| |||| |||  ||||||
Db        240 AGCTGAGCTTTCGCTCACCGTCCACCTCTGCTGTTTGCCGCCATCGCAGACCCACCGCTG 181

Qy        210 ACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCC 269
              ||| |||||||||||||||| ||||| |||| |||||||||||||||||| :||||||
Db        180 ACTTCCATCCCTCTGGATCCAGCAGGATGTCTGCTGTGCTCCTGATCCAGCGGGGCGCCC 121

Qy        270 ATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGG 329
              |||||||||||| ||  |||||||||||||||||||| |||||| |||||||||| |||
Db        120 ATTGCCGCTCCCGATCAGGCTAAAGGCTTGCCATTGTTCCTGCATGGCTAAGTGCCCGGG 61

Qy        330 TTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCA 388
              ||  |  |||| ||||||||| ||| | |||||||||| ||||||||||  || ||||
```

```
Db         60 TTGATCATAATCGAGCTGAACGCTAGTTACTGGGTTCCACGGTTCTCTTCCATGGCCCA 2
```

RESULT 1026
CQ077341
LOCUS CQ077341 275 bp DNA linear PAT 20-JAN-2004
DEFINITION Sequence 13141 from Patent WO0157278.
ACCESSION CQ077341
VERSION CQ077341.1 GI:41047210
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelialcells
JOURNAL Patent: WO 0157278-A 13141 09-AUG-2001; Aeomica, Inc. (US)
FEATURES Location/Qualifiers
source 1. .275
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AE000660.1
EXPRESSED IN HELA, SIGNAL = 1.1
EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
NT HIT: AB024469.1, EVALUE 1.00e-118
SWISSPROT HIT: P31792, EVALUE 3.00e-17"

Query Match 3.3%; Score 244.5; DB 1; Length 275;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 252; Conservative 14; Mismatches 9; Indels 1; Gaps 1;

```
Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||:||:||||||| ||||||:||||| |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              |||||||| ||:||||:|||||:|||||||||||: ||| |||||||||||||||||| .
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||
Db         120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||||:||||:|||||||||||||||||
Db         180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db         240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275
```

RESULT 1027
CQ108356
LOCUS CQ108356 275 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 17215 from Patent WO0157272.
ACCESSION CQ108356
VERSION CQ108356.1 GI:41077409
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for

```
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 17215 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .275
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AE000660.1
                      EXPRESSED IN PLACENTA, SIGNAL = 2.6
                      EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
                      NT HIT: AB024469.1, EVALUE 1.00e-118
                      SWISSPROT HIT: P31792, EVALUE 3.00e-17"

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||:||:|||||| ||||||:||||| |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||| ||:||||:|||||:||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              ||||||||||||||||||||||| |||||:||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||:||||:|||||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 1028
CQ146988
LOCUS       CQ146988                275 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17010 from Patent WO0157276.
ACCESSION   CQ146988
VERSION     CQ146988.1  GI:41154338
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 17010 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .275
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AE000660.1
                      EXPRESSED IN BONE MARROW, SIGNAL = 0.98
                      EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
                      NT HIT: AB024469.1, EVALUE 1.00e-118
                      SWISSPROT HIT: P31792, EVALUE 3.00e-17"

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||:||:|||||| ||||||:||||| |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||| ||:||||:|||||:||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              ||||||||||||||||||||||| |||||:||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||:||||:|||||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 1029
CQ206780
LOCUS       CQ206780                275 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9057 from Patent WO0157271.
ACCESSION   CQ206780
VERSION     CQ206780.1  GI:41192884
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 9057 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .275
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AE000660.1
                      EXPRESSED IN BT474, SIGNAL = 0.99
                      EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
                      NT HIT: AB024469.1, EVALUE 1.00e-118
                      SWISSPROT HIT: P31792, EVALUE 3.00e-17"

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||:||:|||||| ||||||:||||| |||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||| ||:||||:|||||:||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              ||||||||||||||||||||||| |||||:||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:|||||||||||||||||||||||||||||||:||||:|||||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275
```

```
RESULT 1030
CQ230216
LOCUS       CQ230216                275 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17055 from Patent WO0157273.
ACCESSION   CQ230216
VERSION     CQ230216.1  GI:41213434
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 17055 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AE000660.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
                     EST_HUMAN HIT: AU134310.1, EVALUE 1.00e-117
                     NT HIT: AB024469.1, EVALUE 1.00e-118
                     SWISSPROT HIT: P31792, EVALUE 3.00e-17"

  Query Match             3.3%;  Score 244.5;  DB 1;  Length 275;
  Best Local Similarity   91.3%;  Pred. No. 0;
  Matches  252;  Conservative   14;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3800 ACAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTC 3859
              ||||||||||||||||||:||:||||||| ||||||:||||| ||||||||||||||| |||
Db          1 ACAATGAAGAAAAGATAGAACATAACTGCCAACAAGTAATTGCTCAAACCTATGCCGCTC 60

Qy       3860 GAGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGA 3919
              ||||||||  ||:||||:|||||:||||||||||: ||| ||||||||||||||||||||
Db         61 GAGGGGACTTTTTAGAGGTTCCCTTGACTGATCCCCACC-TCAACTTGTATACTGATGGA 119

Qy       3920 AGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTGGTCAGTGATAATGGA 3979
              |||||||||||||||||||||||| |||||:||||||||||||||||||||||||||||||
Db        120 AGTTCCTTTGTAGAAAAAGGACTTCAAAAAGTGGGGTATGCAGTGGTCAGTGATAATGGA 179

Qy       3980 ATAYTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTYAGCTRGCAGAACTAATAGCC 4039
              |||:||||||||||||||||||||||||||||||||||:||||:|||||||||||||||
Db        180 ATACTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGCTCAGCTGGCAGAACTAATAGCC 239

Qy       4040 YTCAYTKGGGCACTAGAATTAGGAGAAGRAAAAAGG 4075
              :|||:| ||| |||||||| ||||||||:|||||||
Db        240 CTCACTCGGGTACTAGAATCAGGAGAAGGAAAAAGG 275


RESULT 1031
AF315117/c
LOCUS       AF315117                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W8 long terminal repeat, partial
            sequence.
ACCESSION   AF315117
VERSION     AF315117.1  GI:12620378
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W8"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.2%;  Score 243.1;  DB 1;  Length 299;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  266;  Conservative    5;  Mismatches   19;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
              | |||||||||||| |||||||||||||| ||:|||||||||:||||||||| ||||||| |
Db        299 AAACAGGACTAGCAGGATTTCCTAGGCCGATTAAGAATCCCTAAGCCTAGATGGGAAGTT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 GACCACATCCACCTTTAAACACGGGGCTTGCAACTCAGCTCACACCCGACCCATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                  ||||||| | |||||||||||||||||||| ||||||||||||||||||||||||||
Db        179 AGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
              |||:||||||:|||||||||||| |||||||||||||| |||||||||||| |  |||
Db        119 CATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCAGGATATAAACCCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||||||||||| | |||| ||||||||||||||||||||||||||||||||||||||||
Db          59 GAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 1


RESULT 1032
AF315115/c
LOCUS       AF315115                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W5 long terminal repeat, partial
            sequence.
ACCESSION   AF315115
VERSION     AF315115.1  GI:12620376
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
```

```
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W5"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.2%;  Score 242.7;  DB 1;  Length 299;
  Best Local Similarity   88.6%;  Pred. No. 0;
  Matches  265;  Conservative    6;  Mismatches   19;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||  ||||||||||||||  |:||||||| :|||||||||:|||||||||
Db        299 AGACAGGACTACCTGGATTTCCTAGGCCAACTAAGAATCTCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             |||||||||||||||||||||| ||||||||||||||:|||||||| |||| |||
Db        239 GACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCTCACACCCGACCCATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 ||||||| | |||||||||||||||||||| |||||||||||||||||| ||||
Db        179 AGAAAGAGAGCCCGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGTCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||||||:|||||||||||| |||||||||||||| |||||||||||||  |  |||
Db        119 CATCTATTGCCTGAGAGCACAGCGGGAGGGACAATGATCAGGATATAAACCCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
             |||||||||||| | |||| ||||||||||||||||||||||||||||||||||||||
Db          59 GAGCCGGCAACGACTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 1


RESULT 1033
BD003154
LOCUS       BD003154                279 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Secretory proteins and polynucleotides encoding the same.
ACCESSION   BD003154
VERSION     BD003154.1  GI:18631115
KEYWORDS    JP 2001501813-A/34.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Mccoy,J.M., Levallie,E.R., Racie,L.A., Merberg,D., Treacy,M.,
            Evans,C. and Spaulding,V.
  TITLE     Secretory proteins and polynucleotides encoding the same
  JOURNAL   Patent: JP 2001501813-A 34 13-FEB-2001;
            GENETICS INSTITUTE INC,KENNETH JACOBS
COMMENT     OS   Unidentified
            PN   JP 2001501813-A/34
            PD   13-FEB-2001
            PF   23-JUL-1997 JP 1998508939
            PR   26-JUL-1996 US    08/686878
            PI   JOHN M MCCOY,EDWARD R LEVALLIE,LISA A RACIE,DAVID MERBERG, PI
             MAURICE TREACY,
            PI   CHERYL EVANS,VIKKI SPAULDING
            PC   C12N15/09,A61K38/00,A61K48/00,A61P3/02,A61P7/02,A61P7/08, PC
            A61P35/00,
            PC   A61P37/00,C07K14/47,C12N5/10,C12P21/02,C12N15/00,C12N5/00, PC
             A61K37/02
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .279
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .279
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
              :|||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 1034
I80073
LOCUS       I80073                  279 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 50 from patent US 5708157.
ACCESSION   I80073
VERSION     I80073.1  GI:3208363
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Evans,C. and Spaulding,V.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5708157-A 50 13-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .279
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             3.2%;  Score 238;  DB 1;  Length 279;
  Best Local Similarity   90.8%;  Pred. No. 0;
  Matches  238;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       7321 ACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380
             :|||||||||||||||||||||||||| ||||||   || | |||| ||||||||||||||
Db          1 RCCACATCCACCTTTAAACACGGGGNTTGCAAANAAGATNACACTTGACCAATCAGAGAG 60

Qy       7381 CTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGC 7440
              ||| ||||||| | ||| |||||| ||||||||||||||||||||||||:||||||
Db         61 NTCANTAAAATGATNATTNGGCAAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGC 120

Qy       7441 MTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAA 7500
              :|||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        121 CTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTNGAGCCGGCAA 180

Qy       7501 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
```

```
Db        181 CGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTNTGTTTTCATGCTATTTCAN 240

Qy       7561 TCTATTAAATCTTGCARCTGCR 7582
              | |||||||| |||||:||||:
Db        241 TNTATTAAATNTTGCAACTGCA 262


RESULT 1035
AF315111/c
LOCUS       AF315111                 299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W2 long terminal repeat, partial
            sequence.
ACCESSION   AF315111
VERSION     AF315111.1  GI:12620372
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W2"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.2%;  Score 237.9;  DB 1;  Length 299;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  262;  Conservative    6;  Mismatches   22;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
              |||||||||||||||||||||||||  |:|||||||||:|||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
              ||| ||| |||||||||| |||  |||||||||||||:|||||| ||||||||
Db        239 GACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                  ||||||||  |||||||||||||||||||||| |||||||||| ||||||||||||||
Db        179 GTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAGAGAAATAGCCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
              |||:||| ||:|||||||||||||||||||||||||||| ||||||||||||||  |||
Db        119 CATCTATCGCCTGAGAGCACAGCAGGAGGGACAATGATCCGGATATAAACCCAAGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              ||||| |||| ||| |||| ||||| |||||||||||||||||||||||||||||| ||||||
Db         59 GAGCCAGCAATGGCTACCCTCTTTGTGTCCCCTCCCTTTGTATGGGAGCTCTATTTTCA 1


RESULT 1036
BD136194
```

```
LOCUS       BD136194                 310 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION   BD136194
VERSION     BD136194.1  GI:23231139
KEYWORDS    JP 2002509437-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
            Perron,H. and Mandrand,B.
  TITLE     Retroviral nucleic material and nucleotide fragments, in
            particular, associated with multiple sclerosis and/or rheumatoid
            arthritis, for diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: JP 2002509437-A 4 26-MAR-2002;
            BIO MERIEUX
COMMENT     OS   Unidentified
            PN   JP 2002509437-A/4
            PD   26-MAR-2002
            PF   07-JUL-1998 JP 1999508255
            PR   07-JUL-1997 FR     97/08816
            PI   GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC  PI
              BEDIN,
            PI   MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,
            PI   BERNARD MANDRAND
            PC   C12N15/48,C12Q1/70,C07K14/15,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Retroviral nucleic material and nucleotide fragments, in  CC
                 particular,
            CC   associated with multiple sclerosis and/or
            rheumatoid arthritis,
            CC            for
            CC   diagnostic, prophylactic and therapeutic uses FH   Key
              Location/Qualifiers
            FT   source          1. .310
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
              ||||||  ||   ||| |||||| |||||||||| || |||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
              ||||||| ||| ||||| || |||||||| |||     | ||| ||||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
              ||||||||| ||||| ||||||||||  |||||| ||| ||||  |||||||||||||||
Db         121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| ||||||||||||||:|||||||||| ||||||
Db         181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db         241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db         301 AAGAAATAAT 310
```

```
RESULT 1037
CS248636
LOCUS       CS248636                310 bp    DNA     linear   PAT 09-JAN-2006
DEFINITION  Sequence 4 from Patent EP1612270.
ACCESSION   CS248636
VERSION     CS248636.1  GI:84662033
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Parahnos-Baccala,G.
  TITLE     Retroviral nucleic material and nucleotide fragments, in particular
            associated with multiple sclerosis and/or rheumatoid arthritis, for
            diagnostic, prophylactic and therapeutic uses
  JOURNAL   Patent: EP 1612270-A 4 04-JAN-2006;
            BIOMERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             |||||| ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||     | ||| |||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||| |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             |||||| |||||||| |||||||||| ||||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 1038
AR344384
LOCUS       AR344384                310 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 100 from patent US 6582703.
ACCESSION   AR344384
VERSION     AR344384.1  GI:33740325
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 100 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             |||||| ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||     | ||| |||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||| |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180

Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
             |||||| |||||||| |||||||||| ||||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
             |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 1039
AX001022
LOCUS       AX001022                310 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 4 from Patent WO9902666.
ACCESSION   AX001022
VERSION     AX001022.1  GI:7241261
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Ott,C. and Bedin,F.
  TITLE     RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR
            ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR
            DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
  JOURNAL   Patent: WO 9902666-A 4 21-JAN-1999;
            BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             3.2%;  Score 237.1;  DB 1;  Length 310;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  266;  Conservative    6;  Mismatches   35;  Indels    3;  Gaps    1;

Qy       4348 GCTTATTAAACTTCCCTTAGTATAGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTC 4407
             |||||| ||   ||| |||||| |||||||||| || ||||||||||||||||||||||
Db          1 GCTTATAGAAGGACCCCTAGTATGGGGTAATCCCCTCTGGGAAACCAAGCCCCAGTACTC 60

Qy       4408 AGCAGGAGAAACAGAATGGGGAACCTCACGAGG---CAGTTTTCTCCCCTCGGGACGGTT 4464
             ||||||| ||| ||||| || |||||||| |||     | ||| |||||||  || || |
Db         61 AGCAGGAAAAATAGAATAGGAAACCTCACAAGGACATACTTTCCTCCCCTCCAGATGGCT 120

Qy       4465 AGCCACTGAAGAAGGGAAAATACTTTTGCCTGCAACTATCCAATGGAAATTACTTAAAAC 4524
             ||||||||| ||||| |||||||||| |||||| ||| ||||  ||||||||||||||||
Db        121 AGCCACTGAGGAAGGAAAAATACTTTCACCTGCAGCTAACCAACAGAAATTACTTAAAAC 180
```

```
Qy       4525 CCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCATTATT 4584
              |||||| |||||||| |||||||||| ||||||||||||||:|||||||||| ||||||
Db        181 CCTTCACCAAACCTTCCACTTAGGCATTGATAGCACCCATCAGATGGCCAAATTATTATT 240

Qy       4585 TACTGGACCAGGCCTTTTCAAAACTATCAAGCARATAKTCAGGGCCTGTGAAKTGTGCCA 4644
              |||||||||||||||||||||||||||||| |:|||:|||||| |||||||:|||||||
Db        241 TACTGGACCAGGCCTTTTCAAAACTATCAAGAAGATAGTCAGGGGCTGTGAAGTGTGCCA 300

Qy       4645 RARAAATAAT 4654
              :|:|||||||
Db        301 AAGAAATAAT 310


RESULT 1040
CQ686782
LOCUS       CQ686782                290 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31708 from Patent WO02070737.
ACCESSION   CQ686782
VERSION     CQ686782.1  GI:42216459
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 31708 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .290
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             3.1%;  Score 235.6;  DB 1;  Length 290;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  256;  Conservative    6;  Mismatches   15;  Indels   10;  Gaps    1;

Qy       7295 AATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACT 7354
              |||||:|||||||||:|||||||||||||| | ||||||||||||||||||||||||||||
Db          14 AATCCCTAAGCCTAGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACT 73

Qy       7355 TAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGG 7414
              |||:||||||| |||||||||||||||||||||||||||||||| ||||| ||||||||
Db          74 TAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTAATTAAGCAAAAACAGGAGG 133

Qy       7415 TAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGAT 7474
              |||||||||||||| ||||:||| ||:|||||||||||||| ||||||||||||||||||||
Db         134 TAAAGAAATAGCCAGTCATCTATCGCCTGAGAGCACAGCAAGAGGGACAATGATCGGGAT 193

Qy       7475 ATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATG 7534
              ||||||||| |  |||||||| |||||||| |||| ||||||||||||||||||||||||
Db         194 ATAAACCCAGGCATTCGAGCCAGCAACGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATG 253

Qy       7535 GGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
              ||||| ||||          |||||||||||||||||||||:||||
Db         254 GGAGCCCTGT----------TTTCACTCTATTAAATCTTGCAACTGC 290


RESULT 1041
BD268836
LOCUS       BD268836                322 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Retrovirus expression vector based on HERVLTR sequence.
ACCESSION   BD268836
VERSION     BD268836.1  GI:33078604
KEYWORDS    JP 2002537846-A/15.
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 322)
  AUTHORS   Moesch,C.L., Schoen,U., Baust,C. and Sarah,R.M.
  TITLE     Retrovirus expression vector based on HERVLTR sequence
  JOURNAL   Patent: JP 2002537846-A 15 12-NOV-2002;
            GSF FORSCHUNGSZENTRUM FUER UMWELT UND GESUNDHEIT GMBH, AUSTRIAN
            NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT
COMMENT     OS   Human endogenous retrovirus
            PN   JP 2002537846-A/15
            PD   12-NOV-2002
            PF   09-MAR-2000 JP 2000603410
            PR   10-MAR-1999 DE 199 10 650.9
            PI   CHRISTINE LEIB MOESCH,ULRIKE SCHOEN,CORINNA BAUST,ROBERT  PI
            MICHAEL SARAH
            PC   C12N15/09,A61K35/76,A61K48/00,A61P31/12,A61P35/00,A61P43/00,
            PC   A61P43/00,
            PC   C12N5/10,C12N7/00//(C12N7/00,C12R1:93),C12N15/00,C12N5/00 CC
            Retrovirus expression vector based on HERVLTR sequence FH   Key
                 Location/Qualifiers
            FT   source          1. .322
            FT                   /organism='Human endogenous retrovirus'.
FEATURES             Location/Qualifiers
     source          1. .322
                     /organism="Human endogenous retrovirus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11827"

  Query Match             3.1%;  Score 234.4;  DB 1;  Length 322;
  Best Local Similarity   80.6%;  Pred. No. 0;
  Matches  283;  Conservative    7;  Mismatches   23;  Indels   38;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
              |||||||||||||||||||||||||||||| |||||||||||||| ||:|||||||||
Db          1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
              :|||||||||:|||||||||||||||||||||||||||||| |||||||||||||||:|
Db         61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
              ||||| |||||| ||         ||||||||||||||||||||||||||||| || ||||
Db        121 CACACTTGACCAGTCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
              |||||||||||||| |||||||:||| |:|||||||||| ||||||||||||||||
Db        181 GAGGTAAAGAAATAGACAATCATCTATCACCTGAGAGCACAGTGGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
              | |||||||||| |  |||||||| ||||| |||||||||||||||  | ||
Db        241 GCATATAAACCCAGGCATTCGAGCCAGCAACAGCAACCCCCTTTGGGAGCTCT------- 293

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
                                  |||||||||||||||||||||||:||||
Db        294 ----------------------GTTTTCACTCTATTAAATCTTGCAACTGC 322


RESULT 1042
AX035744
LOCUS       AX035744                322 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 15 from Patent DE19910650.
ACCESSION   AX035744
VERSION     AX035744.1  GI:11191327
KEYWORDS    .
SOURCE      Human endogenous retrovirus
  ORGANISM  Human endogenous retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1
  AUTHORS   Leib-Moesch,C., Baust,C. and Schoen,U.
  JOURNAL   Patent: DE 19910650-A 15 21-SEP-2000;
            GSF FORSCHUNGSZENTRUM UMWELT (DE)
FEATURES             Location/Qualifiers
     source          1. .322
                     /organism="Human endogenous retrovirus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11827"
```

```
  Query Match             3.1%;  Score 234.4;  DB 1;  Length 322;
  Best Local Similarity   80.6%;  Pred. No. 0;
  Matches  283;  Conservative    7;  Mismatches   23;  Indels   38;  Gaps    2;

Qy       7240 TGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCC 7299
             |||||||||||||||||||||||||||||||| ||||||||||||||| ||:|||||||||
Db             1 TGTTGAGATGGGGGACTGAGAGACAGGACTACCTGGATTTCCTAGGCCGACTAAGAATCC 60

Qy       7300 YTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYT 7359
             :|||||||||:|||||||||||||||||||||||||||||| ||||||||||||||||:|
Db            61 CTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGCT 120

Qy       7360 CACACCTGACCAATC---------AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             ||||| |||||| ||          ||||||||||||||||||||||||||| || ||||
Db          121 CACACTTGACCAGTCAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCTAAAACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCG 7470
             |||||||||||||||| |||||||:|||  |:||||||||||| ||||||||||||||||
Db          181 GAGGTAAAGAAATAGACAATCATCTATCACCTGAGAGCACAGTGGGAGGGACAATGATCG 240

Qy       7471 GGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTG 7530
             | |||||||||| |  |||||||| ||||| ||||||||||||||  | ||
Db          241 GCATATAAACCCAGGCATTCGAGCCAGCAACAGCAACCCCCTTTGGGAGCTCT------- 293

Qy       7531 TATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
                                  ||||||||||||||||||||||:||||
Db          294 ----------------------GTTTTCACTCTATTAAATCTTGCAACTGC 322


RESULT 1043
AF315109/c
LOCUS       AF315109                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W11 long terminal repeat, partial
            sequence.
ACCESSION   AF315109
VERSION     AF315109.1  GI:12620370
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W11"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.1%;  Score 233.1;  DB 1;  Length 299;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  259;  Conservative    6;  Mismatches   25;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||| |||||||||||||||| :|||||||||:|||||||||:||||||  |
Db          299 AGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCCCTAAGCCTAGCTGGGAAATT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             |||||| |||||||||||||||||||||||||||||||| ||||:||||||| ||||||||
Db          239 GACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 || |||||||||||||||||||||||| ||| ||||||||||||||| ||||||||
Db          179 GTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAGGAGGTAAAGAAGTAGCCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:|||||||||| |||||||||||||||||| |||||||||||||  |  |||
Db          119 CATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCAGGATATAAACCCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
             |||| ||   ||| |||| |||||||||||||||||||||||||||||| |||||||||
Db           59 AAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGGAGCCCTGTTTTCA 1


RESULT 1044
AF315114/c
LOCUS       AF315114                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W4 long terminal repeat, partial
            sequence.
ACCESSION   AF315114
VERSION     AF315114.1  GI:12620375
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W4"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.1%;  Score 233.1;  DB 1;  Length 299;
  Best Local Similarity   86.6%;  Pred. No. 0;
  Matches  259;  Conservative    6;  Mismatches   25;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||| |||||||||||||||| :|||||||||:|||||||||:||||||  |
Db          299 AGACAGGACTAGTTGGATTTCCTAGGCTGGCTAAGAATCCCTAAGCCTAGCTGGGAAATT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             |||||| |||||||||||||||||||||||||||||||| ||||:||||||| ||||||||
Db          239 GACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 || |||||||||||||||||||||||| ||| ||||||||||||||| ||||||||
```

```
Db        179 GTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAGGAGGTAAAGAAGTAGCCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
              |||:||| ||:|||||||||| |||||||||||||||| |||||||||||| |  |||
Db        119 CATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCAGGATATAAACCCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| ||   ||| |||| |||||||||||||||||||||||| ||||||||||||
Db         59 AAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGAAGCTCTGTTTTCA 1


RESULT 1045
AF315110/c
LOCUS       AF315110                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W18 long terminal repeat, partial
            sequence.
ACCESSION   AF315110
VERSION     AF315110.1  GI:12620371
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W18"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.1%;  Score 231.5;  DB 1;  Length 299;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  258;  Conservative    6;  Mismatches   26;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
              |||||||||||| |||||||||||| | :|||||||||:|||||||||:||||||  |
Db        299 AGACAGGACTAGTTGGATTTCCTAGGCCGGCTAAGAATCCCTAAGCCTAGCTGGGAAATT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
              |||||| |||||||||||||||||||||||| ||||:||||||| |||||||
Db        239 GACCACGTCCACCTTTAAACACGGGGCTTGCAATTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                  || |||||||||||||||||||||||| ||| ||||||||||||||| ||||||||
Db        179 GTAAAGGGAGCTCACTAAAATGCTAATTAGGGAAAAACAGGAGGTAAAGAAGTAGCCAAT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
              |||:||| ||:|||||||||| |||||||||||||||| |||||||||||| |  |||
Db        119 CATCTATCGCCTGAGAGCACAACAGGAGGGACAATGATCAGGATATAAACCCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| ||   ||| |||| |||||||||||||||||||||||| ||||||||||||
Db         59 AAGCCAGCGGTGGCTACCCTCTTTGGGTCCCCTCCCTTTGTATGGAAGCTCTGTTTTCA 1


RESULT 1046
AY098484
LOCUS       AY098484                288 bp    DNA     linear   PRI 01-MAY-2003
DEFINITION  Macaca mulatta endogenous retrovirus HERV-W/7q nonfunctional
            envelope glycoprotein gene, partial sequence.
ACCESSION   AY098484
VERSION     AY098484.1  GI:30313373
KEYWORDS    .
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 288)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. II. HERV-7q
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
   PUBMED   9835022
REFERENCE   2  (bases 1 to 288)
  AUTHORS   Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
            Rieger,F.
  TITLE     The HERV-W/7q family in the human genome. Potential for protein
            expression and gene regulation
  JOURNAL   Cell. Mol. Biol. 48, 205-212 (2002)
REFERENCE   3  (bases 1 to 288)
  AUTHORS   Alliel,P.M., Lecoeur,L., Goudou,D., Bitoun,M., Rieger,F. and
            Perin,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-2002) U-488, INSERM, 80, rue du General Leclerc,
            Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .288
                     /organism="Macaca mulatta"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9544"
                     /note="endogenous_virus: HERV-W/7q"
     misc_feature    <1. .>288
                     /note="nonfunctional envelope glycoprotein due to
                     mutation; env gene; enverin-like gene"

  Query Match             3.0%;  Score 228.3;  DB 1;  Length 288;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  258;  Conservative   15;  Mismatches   14;  Indels    3;  Gaps    3;

Qy       6374 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db          1 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATYGTTGTTTGAATGGCTCTTCAG 60

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
              |||||||||||||||||||:|||||| :|||||:|||:||| :||||||||||||||||
Db         61 AATCTATGTGCTTCCTCTCRTTCTTA-KGCCCCCTATGACCA-MTACACTGAACAAGATT 118

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTA 6553
              :||||| |:|||||  | |||||||||||||:|||||||||||||||||||||:|:|
Db        119 WATACAATYATGTCGAACCTAAGCCCCGCAAMAAAAGAGTACCCATTCTTCCTTTKGTWA 178

Qy       6554 TAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACC-TCTACT 6612
              |  ||||||| |||||||||| |||||||| :|||:|:|||||||||||||| ||||||
Db        179 TCAGAGCAGGGGTGCTAGGTGGACTAGGTAKKGGCRTKGGCGGTATCACAACCTTCTACT 238

Qy       6613 CAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGT 6662
              |||||||||||||| | ||||||||||:|||||| |||||||||  |||
Db        239 CAGTTCTACTACAAAGTYTCTCAAGAACTRAATGGTGACATGGAANNGGT 288


RESULT 1047
AF315112/c
LOCUS       AF315112                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W22 long terminal repeat, partial
            sequence.
```

```
ACCESSION   AF315112
VERSION     AF315112.1  GI:12620373
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W22"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.0%;  Score 228.3;  DB 1;  Length 299;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  256;  Conservative    6;  Mismatches   28;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             ||||||||||||||||||||||||||||||:||||||||:||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             |||| |||||| |||||||| ||||||||||||||| :||| | ||||||||||
Db        239 GACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACTCATATCTGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 ||||||||  |||||||||||||||||||| |||||||||||||||||||||| |
Db        179 GTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:||| ||||||   || |||||||||||| |||||||||| || |  |||
Db        119 CATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| |||| ||| |||| |||||||||||||||| || ||||||||||||||||||
Db         59 AAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCA 1


RESULT 1048
AF315113/c
LOCUS       AF315113                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W23 long terminal repeat, partial
            sequence.
ACCESSION   AF315113
VERSION     AF315113.1  GI:12620374
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W23"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.0%;  Score 228.3;  DB 1;  Length 299;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  256;  Conservative    6;  Mismatches   28;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             ||||||||||||||||||||||||||||||:||||||||:||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             ||| ||| |||||||||| |||  |||||||||||||:||||||| ||||||||
Db        239 GACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 ||||||||  |||||||||||||||||||| |||||||||||||||||||||| |
Db        179 GTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:||| ||||||   || |||||||||||| |||||||||| || |  |||
Db        119 CATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| |||| ||| |||| |||||||||||||||| || ||||||||||||||||||
Db         59 AAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCA 1


RESULT 1049
AF315107/c
LOCUS       AF315107                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W1 long terminal repeat, partial
            sequence.
ACCESSION   AF315107
VERSION     AF315107.1  GI:12620368
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W1"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.0%;  Score 226.7;  DB 1;  Length 299;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  255;  Conservative    6;  Mismatches   29;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||||||||||||||||||  |:||||||||:||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             ||| ||| ||||||||||||| ||||||||||||||:|||||||  |||||||
Db        239 GACTACACCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCAACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 ||||||||  |||||||||||||||||||||| ||||||||||||||||||||||| |
Db        179 GTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:||| ||||||   || |||||||||||| |||||||||| || |  |||
Db        119 CATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| |||| ||| |||| |||||||||||||||| || ||||||||||||||||||
Db          59 AAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCA 1


RESULT 1050
AF315116/c
LOCUS       AF315116                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W6 long terminal repeat, partial
            sequence.
ACCESSION   AF315116
VERSION     AF315116.1  GI:12620377
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W6"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.0%;  Score 225.1;  DB 1;  Length 299;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  254;  Conservative    6;  Mismatches   30;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||||||||||||||||||  |:||||||||:||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTTCCTAGGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             ||| ||| ||||||||||| ||| ||||||||||||||:|||||||  |||||||
Db        239 GACTACACCCACCTTTAACCACTAGGCTTGCAACTTAGCTCACACCCGACCAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 ||||||||  |||||||||||||||||||||| ||||||||||||||||||||||| |
Db        179 GTAAAGAGAGCTTGCTAAAATGCTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAGT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:||| ||||||   || |||||||||||| |||||||||| || |  |||
Db        119 CATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| |||| ||| |||| |||||||||||||||| || ||||||||||||||||||
Db          59 AAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCA 1


RESULT 1051
AF315108/c
LOCUS       AF315108                299 bp    mRNA    linear   PRI 31-JAN-2001
DEFINITION  Homo sapiens isolate HERV-W10 long terminal repeat, partial
            sequence.
ACCESSION   AF315108
VERSION     AF315108.1  GI:12620369
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Cell type-specific expression and promoter activity of human
            endogenous retroviral long terminal repeats
  JOURNAL   Virology 279 (1), 280-291 (2001)
   PUBMED   11145909
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Schon,U., Seifarth,W., Baust,C., Hohenadl,C., Erfle,V. and
            Leib-Mosch,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Institute of Molecular Virology,
            GSF-National Research Center for Environment and Health,
            Ingoldstaedter Landstrasse 1, Neuherberg D-85764, Germany
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="HERV-W10"
                     /db_xref="taxon:9606"
                     /cell_line="T47D"
                     /tissue_type="mammary carcinoma"
                     /note="endogenous_virus: HERV-W"
     LTR             <1. .>299
                     /note="U3/R region"

  Query Match             3.0%;  Score 223.5;  DB 1;  Length 299;
  Best Local Similarity   84.6%;  Pred. No. 0;
```

```
  Matches  253;  Conservative    6;  Mismatches   31;  Indels    9;  Gaps    1;

Qy       7260 AGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGT 7319
             |||||||||||||||||| ||||||| ||:|||||||||:|||||||||:|||||||||
Db        299 AGACAGGACTAGCTGGATTCCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGT 240

Qy       7320 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATC----- 7374
             |||||||||||||||||||||||||||||||||||||||:||| |||   || ||||
Db        239 GACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCATACCCAACAAATCAGGTA 180

Qy       7375 ----AGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAT 7430
                 |||||||||||||||| || ||||||| || |||||||||| |||| ||||| |
Db        179 GTAAAGAGAGCTCACTAAAATACTGATTAGGCGAAAACAGGAGGTAAGGAAACAGCCAGT 120

Qy       7431 CATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTC 7490
             |||:||| ||:||| ||||||   || |||||||||||| |||||||||| || |  |||
Db        119 CATCTATCGCCTGACAGCACAAGGGGCGGGACAATGATCAGGATATAAACTCAGGCATTC 60

Qy       7491 GAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCA 7549
              |||| |||| ||| |||| |||||||||||||||| || ||||||||||||||||||
Db         59 AAGCCAGCAATGGCTACCCACTTTGGGTCCCCTCCCATTTTATGGGAGCTCTGTTTTCA 1


RESULT 1052
CQ739149
LOCUS       CQ739149                 258 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 25083 from Patent WO02068579.
ACCESSION   CQ739149
VERSION     CQ739149.1  GI:42341443
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 25083 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             3.0%;  Score 221.7;  DB 1;  Length 258;
  Best Local Similarity   92.3%;  Pred. No. 0;
  Matches  241;  Conservative    2;  Mismatches   15;  Indels    3;  Gaps    1;

Qy       6469 ATGRCCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAA 6528
             |||:|||||||||||||||||||||||| ||||||| || ||||||||| ||| |||
Db          1 ATGACCATCTACACTGAACAAGATTTATACAATTATGTCGTACCTAAGCCCCACAAGAAA 60

Qy       6529 AGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGC 6588
             |||||||||||||||||| |||||| |||||||||||||||| | ||||||||||||||
Db         61 AGAGTACCCATTCTTCCTTCTGTTATCGGAGCAGGAGTGCTAGGCGGACTAGGTACTGGC 120

Qy       6589 ATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGG 6648
             |||||| |||||||||||||||||||   |||||||||||||||| ||||||||||
Db        121 ATTGGCAGTATCACAACCTCTACTCTGTT---CTACAAACTATCTCAAAAACTAAATGGT 177

Qy       6649 GACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCA 6708
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        178 GACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTGACTCCCTAGCA 237

Qy       6709 GCAGTAGTCCTTCRAAATCGA 6729
             |||||||||||||:|||| ||
Db        238 GCAGTAGTCCTTCAAAATTGA 258


RESULT 1053
CQ747660
LOCUS       CQ747660                 249 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33594 from Patent WO02068579.
ACCESSION   CQ747660
VERSION     CQ747660.1  GI:42371311
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 33594 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .249
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             2.9%;  Score 218.6;  DB 1;  Length 249;
  Best Local Similarity   92.4%;  Pred. No. 0;
  Matches  230;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1793 ATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTGA 1852
             ||| |||||||||||||||||||||||||||||||||||||||||||| ||||| |||| |
Db          1 ATGGCCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAACCTGGGAAAATCAA 60

Qy       1853 ATGCCTAATAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAAGATTGTCCA 1912
             |||||||||||||||| |||||||| |||||||||||||||| |||||||||||| ||
Db         61 ATGCCTAATAGGGCTTGTTTCCAGTGTGGTCTACAAGGACACTTCAAAAAAGATTGTTCA 120

Qy       1913 AGTAGAAGTAAGCCGCCCCCTTCGTCCATGCCCCTTATTTCAAGGGAATCACTGGAAGGCC 1972
             ||||||| |||||||||| |||||||||||||||| |||| ||||||||||||||||||
Db        121 AGTAGAAATAAGCCGCCCCCTCGTCCATGCCCCTTATGTCAATGGAATCACTGGAAGGCC 180

Qy       1973 CACTGCCCCAGGGGACAAAGGTCTTTTGAGTCAGAAGCCACTAACCAGATGATCCAGCAG 2032
             || || |||| ||||| |||||| | ||||||||||||||||||||||||||||||||||
Db        181 CATTGTCCCAAGGGACGAAGGTCCTCTGAGTCAGAAGCCACTAACCAGATGATCCAGCAG 240

Qy       2033 CAGGACTGA 2041
             |||||||||
Db        241 CAGGACTGA 249


RESULT 1054
AY102609
LOCUS       AY102609                 224 bp    DNA     linear   PRI 17-JUN-2002
DEFINITION  Pan troglodytes HERV-W/7q envelope glycoprotein gene, partial cds.
ACCESSION   AY102609
VERSION     AY102609.1  GI:21435752
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 224)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. II. HERV-7q
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
   PUBMED   9835022
REFERENCE   2  (bases 1 to 224)
  AUTHORS   Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
            Rieger,F.
  TITLE     The HERV-W/7q family in the human genome. Potential for protein
            expression and gene regulation
```

```
  JOURNAL   Cell Mol. Biol. 48, 205-212 (2002)
REFERENCE   3  (bases 1 to 224)
  AUTHORS   Alliel,P.M., Lecoeur,L., Perin,J.P. and Rieger,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2002) U-488, Inserm, 80, rue du General Leclerc,
            Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .224
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /note="endogenous_virus: HERV-W/7q"
     CDS             <1. .>224
                     /codon_start=2
                     /product="envelope glycoprotein"
                     /protein_id="AAM51553.1"
                     /db_xref="GI:21435753"
                     /translation="KPHNKRVPILPFVIGAGVLGGLGTGIGGITTSTQFYYKLSQELN
                     GDMERVADSLVTLQDQLNSLAAVVLQNRRA"

  Query Match             2.9%;  Score 215.6;  DB 1;  Length 224;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches  218;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       6513 TAAGCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGG 6572
             |||||||| |||||||||||||||||||||||||||||||| |||||||||||||||||||
Db             1 TAAGCCCCACAACAAAAGAGTACCCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGG 60

Qy       6573 TGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 6632
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CGGACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATC 120

Qy       6633 TCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 6692
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          121 TCAAGAACTAAATGGTGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCA 180

Qy       6693 ACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAGCTT 6736
             |||||||||||||||||||||||||||||:|||||||||||||||
Db          181 ACTTAACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTT 224


RESULT 1055
A46013
LOCUS       A46013                   297 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 2 from Patent EP0674004.
ACCESSION   A46013
VERSION     A46013.1  GI:2300285
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Mandrand,B., Perron,H., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and pathogenic and/or infections agent MSRV2 linked to
            multiple sclerosis, their nucleic components and their applications
  JOURNAL   Patent: EP 0674004-A 2 27-SEP-1995;
            BIO MERIEUX (FR)
COMMENT     Other publication DE 674004T 960919
            Other publication NO 953925 951204
            Other publication FI 954699 951003
            Other publication AU 1711495 950821
            Other publication CA 2141907 950805
            Other publication FR 2715939 950811
            Other publication FR 2715938 950811
            Other publication FR 2715937 950811
            Other publication FR 2715936 950811.
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||    |  | |  ||||||||  ||||||||||  ||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              |||||||| |||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
              |||||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1056
A57049
LOCUS       A57049                   297 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 2 from Patent WO9628552.
ACCESSION   A57049
VERSION     A57049.1  GI:3713032
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 2 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
              |||    |  | |  ||||||||  ||||||||||  ||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
              |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
              |||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
              |||||||| |||||||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
              |||||||||||||||||:|||||||||||||||||| ||
```

```
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1057
A60114
LOCUS       A60114                   297 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 2 from Patent WO9706260.
ACCESSION   A60114
VERSION     A60114.1  GI:3715130
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 2 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1058
A79464
LOCUS       A79464                   297 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 2 from Patent EP0789077.
ACCESSION   A79464
VERSION     A79464.1  GI:6092472
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 2 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1059
A80198
LOCUS       A80198                   297 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 2 from Patent WO9521256.
ACCESSION   A80198
VERSION     A80198.1  GI:6092930
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H. and Mallet,F.
  TITLE     MSRV1 VIRUS AND MSRV2 PATHOGENIC AND/OR INFECTIOUS AGENT ASSOCIATED
            WITH MULTIPLE SCLEROSIS, NUCLEIC ACID COMPONENTS AND APPLICATIONS
            OF SAME
  JOURNAL   Patent: WO 9521256-A 2 10-AUG-1995;
            BIO MERIEUX (FR); PERRON HERVE (FR)
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
             |||   |  | |  ||||||||  |||||||||| ||||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
             |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
             |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
             ||||||||| |||||||||||||||||||||||||||||||| ||| ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
             |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295
```

```
RESULT 1060
A80307
LOCUS       A80307                    297 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 2 from Patent EP0731168.
ACCESSION   A80307
VERSION     A80307.1  GI:6093034
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 2 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                ||||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1061
AR035350
LOCUS       AR035350                  297 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5871745.
ACCESSION   AR035350
VERSION     AR035350.1  GI:5952018
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Multiple sclerosis related virus
  JOURNAL   Patent: US 5871745-A 2 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1062
AR036191
LOCUS       AR036191                  297 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5871996.
ACCESSION   AR036191
VERSION     AR036191.1  GI:5952859
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Retrovirus agents MSRV1 and MSRV2 associated with multiple
            sclerosis
  JOURNAL   Patent: US 5871996-A 2 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||   |  | |  ||||||||  |||||||||| |||||||||||| ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | |||||||||||||||||| ||||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                ||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||| |||||||||||||||||||||||||||||||| |||  ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||:|||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1063
AR036711
LOCUS       AR036711                  297 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5800980.
ACCESSION   AR036711
VERSION     AR036711.1  GI:5954567
KEYWORDS    .
SOURCE      Unknown.
```

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Detection of MSRV1 virus and MSRV2 pathogen and/or infective agent
            associated with multiple sclerosis, by nucleic acid hybridization
  JOURNAL   Patent: US 5800980-A 2 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1064
AR076973
LOCUS       AR076973                297 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 2 from patent US 5962217.
ACCESSION   AR076973
VERSION     AR076973.1  GI:10003719
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and MSRV2 pathogen and/or infective agent associated
            with multiple sclerosis, and biopolymer constituents thereof
  JOURNAL   Patent: US 5962217-A 2 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1065
AR094694
LOCUS       AR094694                297 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 2 from patent US 6001987.
ACCESSION   AR094694
VERSION     AR094694.1  GI:10021853
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 2 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1066
AR183929
LOCUS       AR183929                297 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 2 from patent US 6342383.
ACCESSION   AR183929
VERSION     AR183929.1  GI:20227898
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Cell lines and viral isolates associated with multiple sclerosis
  JOURNAL   Patent: US 6342383-A 2 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="unassigned DNA"
```

```
  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||    |  | |  ||||||||  |||||||||| |||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||||  |||||||||||||||||||||||||||||||||  |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1067
AR343141
LOCUS       AR343141                297 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 2 from patent US 6579526.
ACCESSION   AR343141
VERSION     AR343141.1  GI:33738643
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 2 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||    |  | |  ||||||||  |||||||||| |||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||||  |||||||||||||||||||||||||||||||||  |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1068
AR344298
LOCUS       AR344298                297 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 2 from patent US 6582703.
ACCESSION   AR344298
VERSION     AR344298.1  GI:33740239
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 2 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             2.9%;  Score 215.2;  DB 1;  Length 297;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  247;  Conservative    6;  Mismatches   24;  Indels    2;  Gaps    2;

Qy       2635 ATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAT 2694
                |||    |  | |  ||||||||  |||||||||| |||||||||||||  ||||||||||
Db         19 ATTTTAGGAGTAAGGAAACCCAACGGACAGTGGA-GGTTAGTGCAAGAACTCAGGATTAT 77

Qy       2695 CAATGGAGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMYTTCCC 2754
                |||| ||||| ||||| | | ||||||||||||||||||| |||||||| |||::|||||
Db         78 CAAT-GAGGCTGTTGTTCCTCTATACCCAGCTGTACCTAACCCTTATACAGTGCTTTCCC 136

Qy       2755 AAATACCAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCCTTCTTCTGCA 2814
                |||||||||||||||||||||||||||||:||||||||||:|||||||||| |||||||
Db        137 AAATACCAGAGGAAGCAGAGTGGTTTACAGTCCTGGACCTTAAGGATGCCTTTTTCTGCA 196

Qy       2815 TCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATACTTCAAACCCARCAT 2874
                |||||||||  |||||||||||||||||||||||||||||||||  |||   ||||||:| |
Db        197 TCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGATCCTTTGAACCCAACGT 256

Qy       2875 CTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGG 2913
                |||||||||||||||||:||||||||||||||||| ||
Db        257 CTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAAGG 295


RESULT 1069
AX806706
LOCUS       AX806706                211 bp    DNA     linear   PAT 25-NOV-2003
DEFINITION  Sequence 286 from Patent WO02090544.
ACCESSION   AX806706
VERSION     AX806706.1  GI:38523198
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Legrain,P., Whiteside,S., Mao,J.I., Khrebtukova,I. and Luo,S.
  TITLE     Protein-protein interactions in adipocyte cells (3)
  JOURNAL   Patent: WO 02090544-A 286 14-NOV-2002;
            Hybrigenics (FR) ; LYNX THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .211
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             2.8%;  Score 207.4;  DB 1;  Length 211;
  Best Local Similarity   95.7%;  Pred. No. 0;
```

```
  Matches  202;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5581 ATGGSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTYACCCTSTTTCACTCTCACT 5640
              ||||:|||||||||||||||||||||||:||:||||:|||:|||||:|||||||||||||
Db          1 ATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACT 60

Qy       5641 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTYACCMAGAGTTTCTATGGAGA 5700
              ||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||
Db         61 GCACCCCCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGA 120

Qy       5701 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC 5760
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        121 ATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 180

Qy       5761 ACCTTCACTGCCCACACCCATATGCCCCGCA 5791
              ||||||||||||||||||||||||||||||
Db        181 ACCTTCACTGCCCACACCCATATGCCCCGCA 211


RESULT 1070
CQ055177
LOCUS       CQ055177                278 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 5997 from Patent WO0157270.
ACCESSION   CQ055177
VERSION     CQ055177.1  GI:41029683
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 5997 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HBL100, SIGNAL = 3.8
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
              ||||||| |||||| |||||| | |||||  | |||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
              |||||||||||||||||| ||||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
              |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
              |  |||||||||||||| |||||||  | ||||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1071
CQ074436
LOCUS       CQ074436                278 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 10236 from Patent WO0157278.
ACCESSION   CQ074436
VERSION     CQ074436.1  GI:41044305
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 10236 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HELA, SIGNAL = 4.7
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
              ||||||| |||||| |||||| | |||||  | |||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
              |||||||||||||||||| ||||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
              |||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
              |  |||||||||||||| |||||||  | ||||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1072
CQ105334
LOCUS       CQ105334                278 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14193 from Patent WO0157272.
ACCESSION   CQ105334
VERSION     CQ105334.1  GI:41074387
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 14193 09-AUG-2001;
            Aeomica, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN PLACENTA, SIGNAL = 4.1
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1073
CQ144028
LOCUS       CQ144028                278 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14050 from Patent WO0157276.
ACCESSION   CQ144028
VERSION     CQ144028.1  GI:41101400
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 14050 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN BONE MARROW, SIGNAL = 6.4
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1074
CQ179508
LOCUS       CQ179508                278 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10904 from Patent WO0157274.
ACCESSION   CQ179508
VERSION     CQ179508.1  GI:41174247
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 10904 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HEART, SIGNAL = 4.5
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1075
CQ203867
```

```
LOCUS       CQ203867                278 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 6144 from Patent WO0157271.
ACCESSION   CQ203867
VERSION     CQ203867.1  GI:41189714
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 6144 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN BT474, SIGNAL = 3.3
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| |||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1076
CQ227235
LOCUS       CQ227235                278 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14074 from Patent WO0157273.
ACCESSION   CQ227235
VERSION     CQ227235.1  GI:41210438
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 14074 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| |||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| |||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1077
CQ265384
LOCUS       CQ265384                278 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13645 from Patent WO0157277.
ACCESSION   CQ265384
VERSION     CQ265384.1  GI:41237972
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 13645 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
                     NT HIT: AF009668.1, EVALUE 4.00e-79
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                     SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;
```

```
Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1078
CQ302470
LOCUS       CQ302470                278 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13575 from Patent WO0186003.
ACCESSION   CQ302470
VERSION     CQ302470.1  GI:41263047
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 13575 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .278
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL008629.9
                      EXPRESSED IN LUNG, SIGNAL = 4.9
                      NT HIT: AF009668.1, EVALUE 4.00e-79
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                      SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1079
CQ339648
LOCUS       CQ339648                278 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13742 from Patent WO0157275.
ACCESSION   CQ339648
VERSION     CQ339648.1  GI:41288719
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 13742 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .278
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL008629.9
                      EXPRESSED IN BRAIN, SIGNAL = 4.9
                      NT HIT: AF009668.1, EVALUE 4.00e-79
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-83
                      SWISSPROT HIT: P31792, EVALUE 1.60e-02"

  Query Match             2.7%;  Score 205;  DB 1;  Length 278;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  252;  Conservative    0;  Mismatches   25;  Indels    4;  Gaps    3;

Qy       2296 ATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCC 2355
             ||||||| |||||| |||||| | |||||  | ||||| |||||||| ||||||||||||
Db          1 ATTATGCCTGAAAGACCCACTCCTTTGTT--GCAGAGAGATTCTAGCCAAAGCAGGGGCC 58

Qy       2356 ATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGG 2415
             |||||||||||||||||| ||||||||||||  ||||||||| || ||||||||||||||
Db         59 ATTATACACCTGAACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGG 117

Qy       2416 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTC 2475
             |||||||||||||||||||||||||||||||||||||||||||||| | |||||||||| ||
Db        118 AATTAATCCTGAAGTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATC 176

Qy       2476 CTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCA 2535
             |  |||||||||||||| |||||||  | ||||||||||||||||||||||| |||| |
Db        177 CCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTA 236

Qy       2536 GACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||||| ||||| | |||||| ||||||||||||||||||||
Db        237 GACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 277


RESULT 1080
BV204004
LOCUS       BV204004                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm213737 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV204004
VERSION     BV204004.1  GI:48173545
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
```

```
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES              Location/Qualifiers
     source           1. .201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>201

  Query Match             2.7%;  Score 199;  DB 1;  Length 201;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  196;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7148 TCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTC 7207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db             1 TCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTC 60

Qy       7208 GGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGA 7267
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db            61 GGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAGAKGGGGGACTGAGAGACAGGA 120

Qy       7268 CTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACAT 7327
             ||||||||||||||||||||||:|||||||||:|||||||||:||||||||||||||||||
Db           121 CTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACAT 180

Qy       7328 CCACCTTTAAACACGGGGCTT 7348
             |||||||||||||||||||||
Db           181 CCACCTTTAAACACGGGGCTT 201


RESULT 1081
BV204007
LOCUS       BV204007                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm213743 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV204007
VERSION     BV204007.1  GI:48173549
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES              Location/Qualifiers
     source           1. .201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>201

  Query Match             2.7%;  Score 199;  DB 1;  Length 201;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  196;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7224 AGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTA 7283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 AGCACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTA 60

Qy       7284 GGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGG 7343
             ||||||:|||||||||:|||||||||:||||||||||||||:||||||||||||||||||
Db            61 GGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCRCATCCACCTTTAAACACGG 120

Qy       7344 GGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 7403
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA 180

Qy       7404 AAGACAGGAGGTAAAGAAATA 7424
             |||||||||||||||||||||
Db           181 AAGACAGGAGGTAAAGAAATA 201


RESULT 1082
BV205765
LOCUS       BV205765                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm219128 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV205765
VERSION     BV205765.1  GI:48175771
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES              Location/Qualifiers
     source           1. .201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>201

  Query Match             2.7%;  Score 199;  DB 1;  Length 201;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  196;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7226 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 7285
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 CACTTAGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGG 60
```

```
Qy       7286 CTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGG 7345
             ||||:|||||||||:|||||||||:||||||||||||||:||||||||||||||||||||
Db         61 CTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACWTCCACCTTTAAACACGGGG 120

Qy       7346 CTTGCAACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 7405
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAA 180

Qy       7406 GACAGGAGGTAAAGAAATAGC 7426
             |||||||||||||||||||||
Db        181 GACAGGAGGTAAAGAAATAGC 201


RESULT 1083
BV205767
LOCUS       BV205767                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm219131 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV205767
VERSION     BV205767.1  GI:48175774
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.7%;  Score 199;  DB 1;  Length 201;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  196;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7231 AGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAY 7290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db          1 AGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAC 60

Qy       7291 TAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGC 7350
             |||||||||:|||||||||:||||||||||||||||||:||||||||||||||||||||||
Db         61 TAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCAYCTTTAAACACGGGGCTTGC 120

Qy       7351 AACTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 7410
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAG 180

Qy       7411 GAGGTAAAGAAATAGCCAATC 7431
             |||||||||||||||||||||
Db        181 GAGGTAAAGAAATAGCCAATC 201


RESULT 1084
BV205773
LOCUS       BV205773                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm219149 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV205773
VERSION     BV205773.1  GI:48175782
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.7%;  Score 199;  DB 1;  Length 201;
  Best Local Similarity   97.5%;  Pred. No. 0;
  Matches  196;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7354 TTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAG 7413
             ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAG 60

Qy       7414 GTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGGA 7473
             |||||||||||||||||||:||||||:||||||||||||:||||||||||||||||||||
Db         61 GTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCRGGAGGGACAATGATCGGGA 120

Qy       7474 TATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTAT 7533
             |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TATAAACCCAAGTTTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTAT 180

Qy       7534 GGGAGCTCTGTTTTCATGCTA 7554
             |||||||||||||||||||||
Db        181 GGGAGCTCTGTTTTCATGCTA 201


RESULT 1085
BV204005
LOCUS       BV204005                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm213739 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV204005
VERSION     BV204005.1  GI:48173546
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
```

```
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.7%;  Score 198.6;  DB 1;  Length 201;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  195;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7187 GAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 7246
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db             1 GAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCAACAGCACTTAGGTTTTCCTGTTGAG 60

Qy       7247 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAYTAAGAATCCYTAAGCC 7306
              ||||||||||||||||||||||||||||||||||||||||:||:|||||||||:||||||
Db            61 ATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCYGACTAAGAATCCCTAAGCC 120

Qy       7307 TAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCT 7366
              |||:||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db           121 TAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCT 180

Qy       7367 GACCAATCAGAGAGCTCACTA 7387
              |||||||||||||||||||||
Db           181 GACCAATCAGAGAGCTCACTA 201


RESULT 1086
G01776
LOCUS       G01776                   240 bp    DNA     linear   STS 06-OCT-1995
DEFINITION  human STS SHGC4-652 clone p4-780, sequence tagged site.
ACCESSION   G01776
VERSION     G01776.1  GI:1009467
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 240)
  AUTHORS   Goold,R.D., diSibio,G.L., Elam,D., Lang,D.B., Maratukulam,A.,
            Perkins,S., Qin,F., Xu,H., Cox,D.R. and Myers,R.M.
  TITLE     STSs from Human Genomic DNA
  JOURNAL   Unpublished (1994)
COMMENT
            Contact: Richard Myers
            Stanford Human Genome Center
            Stanford University School of Medicine
            855 California Avenue, Palo Alto, CA 94304, USA.
            Email: hgmcinfo@camis.stanford.edu

            Primer A: GCTAAAGGCTTACCATTGTTCCTG
            Primer B: TGTTCTCTGACCTGGGGTTCTTG
            STS size: 191
            Protocol:
                      Template:       25 ng
                      Primer:         each 10 pmol
                      dNTPs:          each 200 uM
                      Taq Polymerase: 0.025 units/ul
                      Total Vol:      10 ul
                      Thermocycler:   PE 9600
                      PCR Profile:
                        Initial denaturation: 94 degrees C for 1.5 minutes
                        PCR Cycles:    30
                        Denaturation:         94 degrees C for 15 seconds
                        Annealing:            64 degrees C for 23  seconds
                        Polymerization:       72 degrees C for 3.5 minutes
                        Final extension:      72 degrees C for 3.5 minutes.

            Buffer:
                      MgCl2:          2.5 mM
                      KCl:             50 mM
                      Tris-HCl:        20 mM
                      pH:                8.3.
FEATURES             Location/Qualifiers
     source          1. .240
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="4"
     STS             6. .196
     primer_bind     6. .29
     primer_bind     complement(174. .196)

  Query Match             2.6%;  Score 193.9;  DB 1;  Length 240;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches  214;  Conservative    4;  Mismatches   22;  Indels    1;  Gaps    1;

Qy        283 ATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTG 342
              ||  |||||||||||| ||||||| |||||||||||||||||| ||||| || ||||| |
Db          1 ATCAGGCTAAAGGCTTACCATTGTTCCTGCACGGCTAAGTGCCCGGGTTCGTCCTAATCG 60

Qy        343 AGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKA 402
              |||||||||||| |||||||||||    | |||||||||||||||| ||||||||||:|
Db         61 AGCTGAACACTAGTCACTGGGTTCCCCAGCTCTCTTCTGTGACCCACAGCTTCTAATAGA 120

Qy        403 ACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAACGAA 462
              |||||||||| ||||||||||||||||||||||||||||||||| |||||:||| |||
Db        121 GCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCCATGAGGCCAA-GAA 179

Qy        463 CTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCCTGCTACCRTCTTGG 522
              | ||||||||||||| |||||||||||||||||||||||| ||||||| ||:| ||||
Db        180 CCCCAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGAAGCAGCCTGCTGCCATTTTGG 239

Qy        523 A 523
              |
Db        240 A 240


RESULT 1087
BV205774
LOCUS       BV205774                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm219150 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV205774
VERSION     BV205774.1  GI:48175783
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
```

```
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES              Location/Qualifiers
     source           1. .201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>201

  Query Match              2.5%;  Score 189.6;  DB 1;  Length 201;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches  186;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7391 TGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCAC 7450
             ||||||||||||||||||||||||||||||||||||||||||:||||||:|||||||||
Db          1 TGCTAATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCAC 60

Qy       7451 AGCAGGAGGGACAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCC 7510
             ||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db         61 AGCAGGAGGGACAATGATCGGGATATAAACCCAAGTTTTCRAGCCGGCAACGGCAACCCC 120

Qy       7511 CTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAAT 7570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTTTGGGTCCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAAT 180

Qy       7571 CTTGCARCTGCR 7582
             ||||||:||||:
Db        181 CTTGCAACTGCA 192


RESULT 1088
BV205775
LOCUS       BV205775                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm219151 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV205775
VERSION     BV205775.1  GI:48175785
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES              Location/Qualifiers
     source           1. .201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>201

  Query Match              2.4%;  Score 178.6;  DB 1;  Length 201;
  Best Local Similarity   96.7%;  Pred. No. 0;
  Matches  175;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7402 CAAAGACAGGAGGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGA 7461
             ||||||||||||||||||||||||||||||||||||:|||||:|||||||||||||||||||
Db          1 CAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGA 60

Qy       7462 CAATGATCGGGATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCC 7521
             |||||||||||||||||||||||||:|||||||||||||:|||||||||||||||||||||
Db         61 CAATGATCGGGATATAAACCCAAGTTTTCGAGCCGGCAACRGCAACCCCCTTTGGGTCCC 120

Qy       7522 CTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCARCTGC 7581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        121 CTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGC 180

Qy       7582 R 7582
              :
Db        181 A 181


RESULT 1089
CQ058368/c
LOCUS       CQ058368                200 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 9188 from Patent WO0157270.
ACCESSION   CQ058368
VERSION     CQ058368.1  GI:41032874
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 9188 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .200
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL034410.8
                      EXPRESSED IN HBL100, SIGNAL = 1.9
                      NT HIT: gi11419444, EVALUE 5.00e-78
                      EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                      SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match              2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  |||||||  |||||||||  |||||  |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
```

```
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 1090
CQ077714/c
LOCUS       CQ077714                200 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 13514 from Patent WO0157278.
ACCESSION   CQ077714
VERSION     CQ077714.1  GI:41047583
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 13514 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN HELA, SIGNAL = 1.6
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                     SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 1091
CQ108723/c
LOCUS       CQ108723                200 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17582 from Patent WO0157272.
ACCESSION   CQ108723
VERSION     CQ108723.1  GI:41077776
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 17582 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN PLACENTA, SIGNAL = 6.2
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                     SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 5532
             ||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 1092
CQ147355/c
LOCUS       CQ147355                200 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17377 from Patent WO0157276.
ACCESSION   CQ147355
VERSION     CQ147355.1  GI:41154705
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 17377 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.5
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                     SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
```

```
                ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy        5533 TCTTCAAACAACAACCAGGA 5552
               ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1093
CQ182729/c
LOCUS       CQ182729                200 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14125 from Patent WO0157274.
ACCESSION   CQ182729
VERSION     CQ182729.1  GI:41177515
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 14125 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .200
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL034410.8
                      EXPRESSED IN HEART, SIGNAL = 1.4
                      NT HIT: gi11419444, EVALUE 5.00e-78
                      EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                      SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
               ||||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||
Db         200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy        5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
               |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db         140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy        5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
               |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy        5533 TCTTCAAACAACAACCAGGA 5552
               ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1094
CQ207144/c
LOCUS       CQ207144                200 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9421 from Patent WO0157271.
ACCESSION   CQ207144
VERSION     CQ207144.1  GI:41193300
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 9421 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .200
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL034410.8
                      EXPRESSED IN BT474, SIGNAL = 1.5
                      NT HIT: gi11419444, EVALUE 5.00e-78
                      EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                      SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
               ||||||||||||||||||||||||||||||||||||:||||||||| | ||||||||||
Db         200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy        5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
               |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db         140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy        5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
               |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy        5533 TCTTCAAACAACAACCAGGA 5552
               ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1095
CQ230600/c
LOCUS       CQ230600                200 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17439 from Patent WO0157273.
ACCESSION   CQ230600
VERSION     CQ230600.1  GI:41213818
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 17439 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .200
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL034410.8
                      EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
                      NT HIT: gi11419444, EVALUE 5.00e-78
                      EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                      SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
```

```
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1096
CQ268738/c
LOCUS       CQ268738                200 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 16999 from Patent WO0157277.
ACCESSION   CQ268738
VERSION     CQ268738.1  GI:41241342
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 16999 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                     SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1097
CQ305761/c
LOCUS       CQ305761                200 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 16866 from Patent WO0186003.
ACCESSION   CQ305761
VERSION     CQ305761.1  GI:41266338
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 16866 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN LUNG, SIGNAL = 1.7
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
                     SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db          80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db          20 TCTTCAAGCGACAACCGTGA 1


RESULT 1098
CQ342933/c
LOCUS       CQ342933                200 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 17027 from Patent WO0157275.
ACCESSION   CQ342933
VERSION     CQ342933.1  GI:41292004
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 17027 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL034410.8
                     EXPRESSED IN BRAIN, SIGNAL = 1.7
                     NT HIT: gi11419444, EVALUE 5.00e-78
                     EST_HUMAN HIT: N76991.1, EVALUE 1.00e-63
```

```
                SWISSPROT HIT: P03360, EVALUE 2.00e-06"

  Query Match             2.4%;  Score 177.2;  DB 1;  Length 200;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches  185;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5353 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATYTACCCCAACTGCGGTTAAAGTGG 5412
             ||||||||||||||||||||||||||||||||||||:|||||||| | ||||||||||
Db        200 ATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACCGTGGTTAAAGTGG 141

Qy       5413 CTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAAC 5472
             |||||||||||||||||||||||||||||  ||||||| |||||||||| ||||| |||
Db        140 CTGGAGTGGAGTCTTGGATACATCACACTCAAGTCAAACCCTGGATACTACCAAACGAAA 81

Qy       5473 CTGAAAATCCAGGAGACAACGCTAGCTATTCCTGTGAACCTCTAGAGGATTTGCGCCTGC 5532
             |||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db         80 CTGAAAATCCAGGAGACAACACTAGCTATTCCTGTGAACCTCTAGAGGATCTGCGCCTGC 21

Qy       5533 TCTTCAAACAACAACCAGGA 5552
             ||||||| | ||||||  ||
Db         20 TCTTCAAGCGACAACCGTGA 1


RESULT 1099
G01772
LOCUS       G01772                   217 bp    DNA     linear   STS 06-OCT-1995
DEFINITION  human STS SHGC4-573 clone p4-628, sequence tagged site.
ACCESSION   G01772
VERSION     G01772.1  GI:1009463
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 217)
  AUTHORS   Goold,R.D., diSibio,G.L., Elam,D., Lang,D.B., Maratukulam,A.,
            Perkins,S., Qin,F., Xu,H., Cox,D.R. and Myers,R.M.
  TITLE     STSs from Human Genomic DNA
  JOURNAL   Unpublished (1994)
COMMENT
            Contact: Richard Myers
            Stanford Human Genome Center
            Stanford University School of Medicine
            855 California Avenue, Palo Alto, CA 94304, USA.
            Email: hgmcinfo@camis.stanford.edu

            Primer A: CAATTCAGCAGGAAGCAGTTAGAG
            Primer B: CTGTGTTTAAAGGTGGATGCAGTC
            STS size: 168
            Protocol:
                    Template:       25 ng
                    Primer:         each 10 pmol
                    dNTPs:          each 200 uM
                    Taq Polymerase: 0.025 units/ul
                    Total Vol:      10 ul
                    Thermocycler:   PE 9600
                    PCR Profile:
                      Initial denaturation: 94 degrees C for 1.5 minutes
                      PCR Cycles:    30
                      Denaturation:         94 degrees C for 15 seconds
                      Annealing:            62 degrees C for 23  seconds
                      Polymerization:       72 degrees C for 3.5 minutes
                      Final extension:      72 degrees C for 3.5 minutes.

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:             50 mM
                    Tris-HCl:        20 mM
                    pH:                8.3.
FEATURES             Location/Qualifiers
     source          1. .217
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="4"
     STS             6. .173
     primer_bind     6. .29
     primer_bind     complement(150. .173)

  Query Match             2.3%;  Score 172.2;  DB 1;  Length 217;
  Best Local Similarity   87.1%;  Pred. No. 0;
  Matches  183;  Conservative    4;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       7171 CGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCTCCCCAACAGCACTT 7230
             | ||||||||||||||||||||||||||||  ||: ||||| ||||||||   |||  |||
Db          1 CACCCCAATTCAGCAGGAAGCAGTTAGAGTCGTCATCGGCTAACCTCCCTGGCAGTCCTT 60

Qy       7231 AGGTTTTCCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGAY 7290
              ||||||| |||||||| ||||| |||||||||||||||||||||||||||||| ||||:
Db         61 GGGTTTTCTTGTTGAGACGGGGGCCTGAGAGACAGGACTAGCTGGATTTCCTAGTCTGAC 120

Qy       7291 TAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTGC 7350
             |||||||||:|||||||| |||||||||||  |||||||||||||||| ||||||||
Db        121 TAAGAATCCCTAAGCCTAGNTGGGAAGGTGACTGCATCCACCTTTAAACACAGGGCTTGC 180

Qy       7351 AACTTAGYTCACACCTGACCAATCAGAGAG 7380
             |||||||:|||||||  |||||||||  ||
Db        181 AACTTAGCTCACACCCCACCAATCAGGTAG 210


RESULT 1100
BV173708
LOCUS       BV173708                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm64271 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV173708
VERSION     BV173708.1  GI:48008850
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.2%;  Score 165.8;  DB 1;  Length 201;
  Best Local Similarity   87.6%;  Pred. No. 0;
  Matches  176;  Conservative    3;  Mismatches   22;  Indels    0;  Gaps    0;

Qy        208 TGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGC 267
              ||||  |||||||| | |||| |||||||||||||||||||||||||||| ||:||||
Db          1 TGACCTCCATCCCTCCGCATCCAGCAGGGTGTCCGCTGTGCTCCTGATCCAGTGAGGCGC 60
```

```
Qy        268 CCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCACGGCTAAGTGCCTG 327
              || ||||| ||||| |||||||||| ||||||||| ||| :||||| ||||||||||||||
Db         61 CCGTTGCCACTCCCGATTGGGCTAAGGGCTTGCCAATGTTMCTGCATGGCTAAGTGCCTG 120

Qy        328 GGTTTGTTCTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCC 387
              |||| || ||||| |||||||||||| |||||||||||||||||||||||  ||||||
Db        121 GGTTCGTCCTAATCGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCCATGACCC 180

Qy        388 ACGGCTTCTAATAKAACTATA 408
              ||  |||||||||:| |||||
Db        181 ACAACTTCTAATAGAGCTATA 201


RESULT 1101
G30805
LOCUS       G30805                   198 bp    mRNA    linear   STS 27-OCT-1998
DEFINITION  sWSS3596 Eric D. Green Homo sapiens STS cDNA clone 7H01C09,
            sequence tagged site.
ACCESSION   G30805
VERSION     G30805.1  GI:3800590
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198)
  AUTHORS   Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,
            Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.
  TITLE     2006 expressed-sequence tags derived from human chromosome
            7-enriched cDNA libraries
  JOURNAL   Genome Res. 7 (3), 281-292 (1997)
  PUBMED    9074931
REFERENCE   2  (bases 1 to 198)
  AUTHORS   Green,E.D.
  TITLE     Human chromosome 7 STSs (1997)
  JOURNAL   Unpublished (1997)
COMMENT     On Oct 28, 1998 this sequence version replaced gi:1923091.
            Synonyms: 7H01C09
            GDB: GDB:4585385
            Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
            Tel: 3014020201
            Fax: 3014024735
            Email: egreen@nhgri.nih.gov
            Primer A: CAATATCACTTACACTGCC
            Primer B: CGTTCATTTTCTTGACTTTTCC
            STS size: 61
            PCR Profile:
                    Presoak:          0 degrees C for  0.00 minute(s)
                    Denaturation:    92 degrees C for  0.17 minute(s)
                    Annealing:       55 degrees C for  1.00 minute(s)
                    Polymerization:  72 degrees C for  1.00 minute(s)
                    PCR Cycles:      35
                    Thermal Cycler:  PerkinElmer 9600
            Protocol:
                    Template:        30-100 ng
                    Primer:          each 1 uM
                    dNTPs:           each 200 uM
                    Taq Polymerase:  0.05 units/ul
                    Total Vol:       10 ul

            Buffer:
                    MgCl2:           1.5 mM
                    KCl:             100 mM
                    Tris-HCl:        10 mM
                    NH4Cl:           5 mM
                    pH:              8.6

              The sequence for this STS was derived from a single    sequencing
              read of direct-selected cDNA clone 7H01C09.   For additional
              information about the NHGRI chromosome 7   mapping project, see
              http://www.nhgri.nih.gov/DIR/GTB/CHR7.   Also see Genomics
              11:548-64 (1991) [MUID=92128937].
FEATURES              Location/Qualifiers
     source           1. .198
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /map="7"
                      /clone="7H01C09"
                      /clone_lib="Eric D. Green"
     STS              106. .166
     primer_bind      106. .124
     primer_bind      complement(145. .166)

  Query Match             2.2%;  Score 165.6;  DB 1;  Length 198;
  Best Local Similarity   89.4%;  Pred. No. 0;
  Matches  177;  Conservative    1;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       4855 CATGAAATAATTCCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCT 4914
              |||||||||||||| ||||| ||||||||||||||||||||||||||||| ||||||||
Db          1 CATGAAATAATTCCTAGATTTGGACTTCCCCGAGGCTTACAGAGTGACAATGGCCCTGCT 60

Qy       4915 TTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTTACAC 4974
              ||| ||||  |||||||||||||||| ||||||||| |||||| |||| ||||||||||||
Db         61 TTCAAGGCTGCAGTAACCCAGGGACTATCCCAGGTGTTAGGCATACAATATCACTTACAC 120

Qy       4975 TGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAGGA 5034
              ||| |||| |||||||| |||||||| || ||| |||||||||| ||||:|||||||  |
Db        121 TGCCCCTGGAGGCCACAATCCTCAGGAAAAGTCAAGAAAATGAACGAAACACTCAAACAA 180

Qy       5035 CATCTAAAAAAGCAAACC 5052
              |||||||||||||| ||||
Db        181 CATCTAAAAAAGCTAACC 198


RESULT 1102
CQ746872/c
LOCUS       CQ746872                 222 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 32806 from Patent WO02068579.
ACCESSION   CQ746872
VERSION     CQ746872.1  GI:42368453
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 32806 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES              Location/Qualifiers
     source           1. .222
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"

  Query Match             2.2%;  Score 165.6;  DB 1;  Length 222;
  Best Local Similarity   86.4%;  Pred. No. 0;
  Matches  191;  Conservative    1;  Mismatches   27;  Indels    2;  Gaps    1;

Qy         95 TCACTCTATTAAATCTTGCAACT--GCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGC 152
              |||||||||||| |||||||||  || ||||||||||  ||||| ||||||||||||||
Db        222 TCACTCTATTAAAACTTGCAACTGCGCTCTCTTCTGGTCAGTGTTTGTTACGGCTCGAGC 163

Qy        153 TGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACT 212
              ||||||||  ||  | ||||||||||||||||||| |   | || ||| ||||||||||
Db        162 TGAGCTTTCACTTGCTGTCCACCACTGCTGTTTGCCGCTGTCACAGACCCGCCGCTGACT 103
```

```
Qy        213 CCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATT 272
              ||||||||||||||||| |||||||||| |||||||||||||||||||| |:| |||||||
Db        102 TCCATCCCTCTGGATCCGGCAGGGTGTCTGCTGTGCTCCTGATCCAGCAAGGTGCCCATT 43

Qy        273 GCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              ||||||| | ||  |||||||||||||||||||| ||||||
Db         42 GCCGCTCTCGATCAGGCTAAAGGCTTGCCATTGTTCCTGCA 2


RESULT 1103
BV172221
LOCUS       BV172221                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm49367 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV172221
VERSION     BV172221.1  GI:48006951
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.2%;  Score 162.8;  DB 1;  Length 201;
  Best Local Similarity   89.2%;  Pred. No. 0;
  Matches  181;  Conservative    4;  Mismatches   16;  Indels    2;  Gaps    1;

Qy       7353 CTTAGYTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGA 7412
              |||||:|||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTAGCTCACACCCAACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGA 60

Qy       7413 GGTAAAGAAATAGCCAATCATYTATTGCMTGAGAGCACAGCAGGAGGGACAATGATCGGG 7472
              ||||||||||||||||||||:||||||:||||||||||: |||||||||| |||| ||
Db         61 GGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGYGGGAGGGACAAGGATCAGG 120

Qy       7473 ATATAAACCCAAGTYTTCGAGCCGGCAACGGCAACCCCCTTTGGGTCCCCTCCCTTTGTA 7532
              ||||||| ||| |  ||| |||| |||| ||||||||||||||||| |||||  |||||
Db        121 ATATAAATCCAGGCATTCCAGCCAGCAATGGCAACCCCCTTTGGGTCGCCTCC--TTGTA 178

Qy       7533 TGGGAGCTCTGTTTTCATGCTAT 7555
              |||||||||||||||  ||||
Db        179 TGGGAGCTCTGTTTTCACTCTAT 201


RESULT 1104
BV174185/c
LOCUS       BV174185                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm69901 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV174185
VERSION     BV174185.1  GI:48009446
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.1%;  Score 158.2;  DB 1;  Length 201;
  Best Local Similarity   89.8%;  Pred. No. 0;
  Matches  177;  Conservative    2;  Mismatches   16;  Indels    2;  Gaps    1;

Qy         93 TTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGC 152
              |||||||||||| |||||||||||||||||||||||||||||| ||||| |||| ||| ||||
Db        197 TTTCACTCTATTCAATCTTGCAACTGCACTCTTCTGGTCCGTGTTTGTTACAGCTTGAGC 138

Qy        153 TGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCC--GCTGA 210
              |||||||| |||| || |||||||||||||||||:||| |||| |||||||  |||||
Db        137 TGAGCTTTCGCTCGCCTTCCACCACTGCTGTTTGCCRCCATCGCAGACCTGCCGTGCTGA 78

Qy        211 CTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCA 270
              || |||||||||| |||| |||||||||||||||||||| ||||||||||:|||||||
Db         77 CTTCCATCCCTCTAGATCTGGCAGGGTGTCCGCTGTGCTCTTGATCCAGCGAGGCGCCCA 18

Qy        271 TTGCCGCTCCCAATTGG 287
              |||||||||| |||||
Db         17 TTGCCGCTCCCGATTGG 1


RESULT 1105
CQ745979/c
LOCUS       CQ745979                222 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31913 from Patent WO02068579.
ACCESSION   CQ745979
VERSION     CQ745979.1  GI:42365605
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 31913 06-SEP-2002;
```

```
           PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .222
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             2.1%;  Score 157.6;  DB 1;  Length 222;
  Best Local Similarity   84.2%;  Pred. No. 0;
  Matches  186;  Conservative    1;  Mismatches   32;  Indels    2;  Gaps    1;

Qy         95 TCACTCTATTAAATCTTGCAACTG--CACTCTTCTGGTCCATGTTTCTTACGGCTCGAGC 152
              ||||||||||||||| |||||||||  ||||||||||||  ||||| ||||||||||||||
Db        222 TCACTCTATTAAATATTGCAACTGCACACTCTTCTGGTCTGTGTTTGTTACGGCTCGAGC 163

Qy        153 TGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACT 212
              ||||||||  ||| | ||||||| ||||||||||||||    |||| ||||||| ||||||
Db        162 TGAGCTTTCACTCGCTGTCCACCCCTGCTGTTTGCCACTGTCGCAGACCTGCCACTGACT 103

Qy        213 CCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATT 272
              ||| ||||| |||||| ||||||| || ||||||||||||||||||| |: | ||||||
Db        102 TCCACCCCTCCGGATCCAGCAGGGTCTCTGCTGTGCTCCTGATCCAGCAAGACACCCATT 43

Qy        273 GCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              ||| | |   || |||||| ||||| |||||||| ||||||
Db         42 GCCACACTGGATCGGGCTAGAGGCTCGCCATTGTTCCTGCA 2


RESULT 1106
CQ116725
LOCUS       CQ116725                202 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25584 from Patent WO0157272.
ACCESSION   CQ116725
VERSION     CQ116725.1  GI:41086595
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 25584 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .202
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 4.2
                     NT HIT: AF156961.1, EVALUE 2.00e-89
                     EST_HUMAN HIT: T96421.1, EVALUE 1.00e-45
                     SWISSPROT HIT: Q02457, EVALUE 8.50e+00"

  Query Match             2.1%;  Score 156.6;  DB 1;  Length 202;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  183;  Conservative    5;  Mismatches   14;  Indels    2;  Gaps    2;

Qy        973 TAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCCCAATTATGACCCCTCC 1032
              |||||||| ||||||| | |||||||||||||||||||||||||  ||||| ||||| |
Db          1 TAGACAAAGGGGTAAATAATGAACCAAAGAGTGCCAATATTCCCCGTTTATGCCCCCT-C 59

Qy       1033 CAAGCAGTGGGAGGAAGAGAATTCGGCCCAGCCAGAGTGCATGTGCYTTTTYYTCTCCCA 1092
              |||||||||||||||| |||||||||||||||||||||||||||| |:||||::|||| ||
Db         60 CAAGCAGTGGGAGGAGGAGAATTCGGCCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCA 119

Qy       1093 GACTTAAAGCAAATAAAAACAGACTTAGGTAAATTCTCAGATAAYCCTGATGGCTATATT 1152
              ||||| |||||||| |||| |||| ||||| ||||||||||||:||||||||||||||||
Db        120 GACTTGAAGCAAATTAAAATAGACCTAGGTCAATTCTCAGATAACCCTGATGGCTATATT 179

Qy       1153 GRTGTTTTACAAGGGTTAGGACAA 1176
              |:|| ||||||||||||||||||||
Db        180 GATG-TTTACAAGGGTTAGGACAA 202


RESULT 1107
BV178412
LOCUS       BV178412                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm100386 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV178412
VERSION     BV178412.1  GI:48014759
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201

  Query Match             2.1%;  Score 156.4;  DB 1;  Length 201;
  Best Local Similarity   86.1%;  Pred. No. 0;
  Matches  180;  Conservative    2;  Mismatches   19;  Indels    8;  Gaps    1;

Qy         65 CCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCT 124
              |||||||||||||||| | |        ||||||||||||||||||||||||||||||||
Db          1 CCTTTGTATGGGAGCTCTGT--------TTTCACTCTATTAAATCTTGCAACTGCACTCT 52

Qy        125 TCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTT 184
              |||||||| ||||| ||||||||||||||||||||  ||| || |||:||||||||||||
Db         53 TCTGGTCCGTGTTTGTTACGGCTCGAGCTGAGCTTTCACTCTCCATCCRCCACTGCTGTT 112

Qy        185 TGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCT 244
              |||| ||| ||||  |||||| |||||| ||||||||||||||||  |||||||||||| |
Db        113 TGCCGCCATCGCAGGCCTGCCACTGACTTCCATCCCTCTGGATCTAGCAGGGTGTCCGTT 172

Qy        245 GTGCTCCTGATCCAGCGARGCGCCCATTG 273
              |||||||||||||||| ||: |||||||||
Db        173 GTGCTCCTGATCCAGTGAGACGCCCATTG 201


RESULT 1108
CQ106174
LOCUS       CQ106174                207 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 15033 from Patent WO0157272.
ACCESSION   CQ106174
VERSION     CQ106174.1  GI:41075227
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 15033 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .207
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN PLACENTA, SIGNAL = 3.3
                     EST_HUMAN HIT: AA368311.1, EVALUE 4.00e-63
                     SWISSPROT HIT: P31792, EVALUE 7.00e-03
                     NT HIT: AB026898.1, EVALUE 1.00e-65"

  Query Match             2.1%;  Score 156.2;  DB 1;  Length 207;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  188;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       2369 ACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAA 2428
              ||||| |||||||||||  |||||||| || ||||||||||||||||||||||||||||
Db          1 ACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGGAATTAATCCTGAA 59

Qy       2429 GTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAA 2488
              |||||||||||||||||||||||||||||| | |||||||||| |||  ||||||||||
Db         60 GTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATCCCATTCAAGTTAA 118

Qy       2489 ACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCA 2548
              |||| |||||||  | |||||||||||||||||||||||| |||| |||||| ||||| |
Db        119 ACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTAGACCCGAGGCCGA 178

Qy       2549 ACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              |||||| ||||||||||||||||||||
Db        179 TCAAGGACTCCAAAAGATTGTTAAGGAC 206


RESULT 1109
CQ228049
LOCUS       CQ228049                207 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14888 from Patent WO0157273.
ACCESSION   CQ228049
VERSION     CQ228049.1  GI:41211267
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 14888 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .207
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
                     EST_HUMAN HIT: AA368311.1, EVALUE 4.00e-63
                     SWISSPROT HIT: P31792, EVALUE 7.00e-03
                     NT HIT: AB026898.1, EVALUE 1.00e-65"

  Query Match             2.1%;  Score 156.2;  DB 1;  Length 207;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  188;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    2;

Qy       2369 ACATAGGAGAAGGAACACCCGTTTGTTGTNCCCCTGCTTGAGGAAGGAATTAATCCTGAA 2428
              ||||| |||||||||||  |||||||| || ||||||||||||||||||||||||||||
Db          1 ACATAAGAGAAGGAACACAAGTTTGTTGT-CCTCTGCTTGAGGAAGGAATTAATCCTGAA 59

Qy       2429 GTCTGGGCAACAGAAGGACAATATGGACGAGCCAAAGAATGCCCGTCCTGTTCAAGTTAA 2488
              |||||||||||||||||||||||||||||| | |||||||||| |||  ||||||||||
Db         60 GTCTGGGCAACAGAAGGACAATATGGACGAG-CGAAGAATGCCCATCCCATTCAAGTTAA 118

Qy       2489 ACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCCCCTCAGACCCAAGGCCCA 2548
              |||| |||||||  | |||||||||||||||||||||||| |||| |||||| ||||| |
Db        119 ACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTCCCTTAGACCCGAGGCCGA 178

Qy       2549 ACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              |||||| ||||||||||||||||||||
Db        179 TCAAGGACTCCAAAAGATTGTTAAGGAC 206


RESULT 1110
AY098483
LOCUS       AY098483                190 bp    DNA     linear   PRI 22-MAY-2002
DEFINITION  Homo sapiens enverin-like protein gene, partial sequence.
ACCESSION   AY098483
VERSION     AY098483.1  GI:21103958
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 190)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. II. HERV-7q
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
   PUBMED   9835022
REFERENCE   2  (bases 1 to 190)
  AUTHORS   Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
            Rieger,F.
  TITLE     The HERV-W/7q family in the human genome. Potential for protein
            expression and gene regulation
  JOURNAL   Cell Mol. Biol. 48, 205-212 (2002)
REFERENCE   3  (bases 1 to 190)
  AUTHORS   Alliel,P.M., Perin,J.P., Lecoeur,L. and Rieger,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-2002) U488, INSERM, 80, rue du General Leclerc,
            Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .190
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     misc_feature    <1. .>190
                     /note="enverin-like protein; similar to endogenous
                     retrovirus HERV-W/7q putative envelope glycoprotein"

  Query Match             2.0%;  Score 148.9;  DB 1;  Length 190;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  163;  Conservative    5;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       6366 CTGCCTACCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGG 6425
              |||| |||||||||||||| |||||||||||||||||||||||||:|||||||||||||
Db          1 CTGCNTACCCTCAGGAATATCTTTTGTCTGTGGTACCTCAGCCTATCRTTGTTTGAATGG 60

Qy       6426 CTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGA 6485
              ||||||||||||||||||||||||||||||||||||||||:|:|:||:||||||| ||
```

```
Db         61 CTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTAKGACCWTCTACACRGA 120

Qy       6486 ACAAGATTTATACAGTTATGT-CATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTC 6544
              |||| ||| ||||| | |||| | || ||||||||||||  |||    |||| | ||||
Db        121 ACAANATTAATACAATCATGTCCNTACCTAAGCCCCGCACAAAACNNCTACCCNTNCTTC 180

Qy       6545 CTTTTGTTAT 6554
              |   |||||
Db        181 NTNNGGTTAT 190


RESULT 1111
AY098485
LOCUS       AY098485                 165 bp    DNA     linear   PRI 19-MAY-2002
DEFINITION  Pan troglodytes endogenous retrovirus HERV-W/7q envelope
            glycoprotein (env) gene, partial cds.
ACCESSION   AY098485
VERSION     AY098485.1  GI:20977649
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 165)
  AUTHORS   Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
            Rieger,F.
  TITLE     Endogenous retroviruses and multiple sclerosis. II. HERV-7q
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
   PUBMED   9835022
REFERENCE   2  (bases 1 to 165)
  AUTHORS   Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
            Rieger,F.
  TITLE     The HERV-W/7q family in the human genome. Potential for protein
            expression and gene regulation
  JOURNAL   Cell. Mol. Biol. 48, 205-212 (2002)
REFERENCE   3  (bases 1 to 165)
  AUTHORS   Alliel,P.M., Lecoeur,L., Rieger,F. and Perin,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-2002) U-488, INSERM, 80, rue du General Leclerc,
            Le Kremlin-Bicetre 94270, France
FEATURES             Location/Qualifiers
     source          1. .165
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /note="endogenous_virus: HERV-W/7q"
     gene            <1. .>165
                     /gene="env"
     CDS             <1. .>165
                     /gene="env"
                     /note="enverin-like protein"
                     /codon_start=3
                     /product="envelope glycoprotein"
                     /protein_id="AAM28225.1"
                     /db_xref="GI:20977650"
                     /translation="SGIFFVCGTSAYXCLNGSSESMCFLSFLVPPXTXXXXQBXYNXV
                     XXKPRXKXXX"

  Query Match             1.9%;  Score 139.6;  DB 1;  Length 165;
  Best Local Similarity   81.9%;  Pred. No. 0;
  Matches  127;  Conservative   24;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6374 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAG 6433
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db          1 CCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCRTTGTTTGAATGGCTCTTCAG 60

Qy       6434 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCYATGRCCATCTACACTGAACAAGATT 6493
              ||||||||||||||||||||||||||||||||:|:::||:||:|:|::||:||::|||
Db         61 AATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCWAKRACCWTCWAMAYKGAMCARRATT 120

Qy       6494 TATACAGTTATGTCATATCTAAGCCCCGCAACAAA 6528
              ::|||| | ::|| ::| ::|||||||:|::|||
Db        121 WWTACAATAWKGTNRWACYYAAGCCCCGMAMMAAA 155


RESULT 1112
A60152
LOCUS       A60152                   150 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 40 from Patent WO9706260.
ACCESSION   A60152
VERSION     A60152.1  GI:3715167
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 40 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
              |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  ||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 1113
A79502
LOCUS       A79502                   150 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 40 from Patent EP0789077.
ACCESSION   A79502
VERSION     A79502.1  GI:6092509
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 40 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
              |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
```

```
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  ||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 1114
AR344331
LOCUS       AR344331                150 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 36 from patent US 6582703.
ACCESSION   AR344331
VERSION     AR344331.1  GI:33740272
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 36 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             1.9%;  Score 139.2;  DB 1;  Length 150;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  141;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2799 GATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 2858
              |||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAGAT 60

Qy       2859 ACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGY 2918
              |||  ||||||:| ||||||||||||||||||:|||||||||||||||||||||||||:
Db         61 CCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGATAGC 120

Qy       2919 CCCCATCTATTTGGCCAGGCATTAGCCCAA 2948
              |||||||||||||||||||||||||||||
Db        121 CCCCATCTATTTGGCCAGGCATTAGCCCAA 150


RESULT 1115
CQ743973/c
LOCUS       CQ743973                174 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 29907 from Patent WO02068579.
ACCESSION   CQ743973
VERSION     CQ743973.1  GI:42358338
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 29907 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .174
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.8%;  Score 137.7;  DB 1;  Length 174;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches  156;  Conservative    1;  Mismatches   16;  Indels    1;  Gaps    1;

Qy        140 TTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANA 199
              ||||||||| |||||||||||||||| |||||||||||||||||||||| ||  |||| |
Db        174 TTACGGCTCAAGCTGAGCTTTTGCTCGCCGTCCACCACTGCTGTTTGCCGCCGTCGCAGA 115

Qy        200 CCTGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAG 259
              || |||||||||| |||| |||||||||| |||||||||| ||||||||||||||||||
Db        114 CCCGCCGCTGACTTCCAT-CCTCTGGATCCAGCAGGGTGTCTGCTGTGCTCCTGATCCAG 56

Qy        260 CGARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              ||:||||||||||| |||| || ||||||||||||||||||||| ||||||
Db         55 TGAGGCGCCCATTGCCACTCCGGATCGGGCTAAAGGCTTGCCATTGTTCCTGCA 2


RESULT 1116
CQ739060/c
LOCUS       CQ739060                180 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 24994 from Patent WO02068579.
ACCESSION   CQ739060
VERSION     CQ739060.1  GI:42341099
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 24994 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .180
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.8%;  Score 136.9;  DB 1;  Length 180;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches  154;  Conservative    3;  Mismatches   15;  Indels    1;  Gaps    1;

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              |||||  | |||||||||| |||||||||||| |||||||||| ||||||||||||||||
Db        180 CTAATCAAACTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTC 121

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGS 455
              |||||:| |||||||||| ||  |||||||||||||||||||||||||||||||||||:
Db        120 TAATAGAGCTATAACACTCACTGCATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGC 61

Qy        456 CAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||| |||| |||||||||||| ||||:|||| |||||||||||||| | ||
Db         60 CAA-GAACCCCAGGTCAGAGAACACGAGGCTTACCACCATCTTGGAAGTGCCC 9


RESULT 1117
CQ752437/c
LOCUS       CQ752437                186 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 38371 from Patent WO02068579.
ACCESSION   CQ752437
VERSION     CQ752437.1  GI:42388618
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 38371 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .186
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.8%;  Score 132.8;  DB 1;  Length 186;
  Best Local Similarity   84.9%;  Pred. No. 0;
  Matches  146;  Conservative    1;  Mismatches   25;  Indels    0;  Gaps    0;

Qy        142 ACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACC 201
              || |||||||||||| |||| ||| || ||| ||||||||||||||| ||   ||| |||
Db        173 ACCGCTCGAGCTGATCTTTCGCTTTCCATCCCCCACTGCTGTTTGCCGCCGTTGCAGACC 114

Qy        202 TGCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCG 261
              |||||||||||| |||||| ||||||||| ||||| |||| |||||||||||||||||| |
Db        113 TGCCGCTGACTTCCATCCTTCTGGATCCAGCAGGTTGTCTGCTGTGCTCCTGATCCAGTG 54

Qy        262 ARGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              |:|| ||||||||| |||| ||||||||||||||||||| |||| || |||
Db         53 AGGCCCCCATTGCCACTCCTGATTGGGCTAAAGGCTTGCCTTTGTTCCCGCA 2


RESULT 1118
CQ742464/c
LOCUS       CQ742464                180 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28398 from Patent WO02068579.
ACCESSION   CQ742464
VERSION     CQ742464.1  GI:42353137
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 28398 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .180
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.7%;  Score 127.3;  DB 1;  Length 180;
  Best Local Similarity   85.5%;  Pred. No. 0;
  Matches  148;  Conservative    3;  Mismatches   21;  Indels    1;  Gaps    1;

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              ||||| |||||||||||| || |||||||||| |||||||||| ||||||||||||||
Db        180 CTAATCGAGCTGAACACTAGTCGCTGGGTTCCACGGTTCTCTTCCGTGACCCACGGCTTC 121

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGS 455
              |||||:| |||||||||| ||  ||||||||| | |||||||| ||||||||| |||||:
Db        120 TAATAGAGCTATAACACTCACTGCATGGCCCAGGGTTCCATTCTTTGGAATCCATGAGGC 61

Qy        456 CAACGAACTCCAGGTCAGAGAATACGARGCTTGCCACCATCTTGGAAGCGGCC 508
              ||| |||| |||||||||||||| |  |:|||||  |||||||| || ||||
Db         60 CAA-GAACCCCAGGTCAGAGAACAAAAGGCTTGCTGCCATCTTGGGAGTGGCC 9


RESULT 1119
CQ740255/c
LOCUS       CQ740255                172 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 26189 from Patent WO02068579.
ACCESSION   CQ740255
VERSION     CQ740255.1  GI:42345404
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 26189 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .172
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.7%;  Score 126.7;  DB 1;  Length 172;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches  148;  Conservative    1;  Mismatches   21;  Indels    1;  Gaps    1;

Qy        144 GGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTG-TTTGCCACCACCGCANACCT 202
              ||||| |||||||||||||||| | |||||||||||||| |||||| ||  |||| |||
Db        172 GGCTCAAGCTGAGCTTTTGCTCGCTGTCCACCACTGCTGTTTTGCCGCCGTCGCAGACCC 113

Qy        203 GCCGCTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGA 262
              ||||||| ||||||||| |||||| ||| |||||| ||||||||||||||||||||||
Db        112 ATTGCTGACTTCCATCCCTCCGGATCCAGCAAGGTGTCTGCTGTGCTCCTGATCCAGCGA 53

Qy        263 RGCGCCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              :|| |||||||||||||| ||  |||||||||||||||||||| ||||||
Db         52 GGCACCCATTGCCGCTCCCGATCAGGCTAAAGGCTTGCCATTGTTCCTGCA 2


RESULT 1120
CQ058491
LOCUS       CQ058491                152 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 9311 from Patent WO0157270.
ACCESSION   CQ058491
VERSION     CQ058491.1  GI:41032997
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 9311 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN HBL100, SIGNAL = 1.1
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
```

```
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 1121
CQ077834
LOCUS       CQ077834                152 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 13634 from Patent WO0157278.
ACCESSION   CQ077834
VERSION     CQ077834.1  GI:41047703
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 13634 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN HELA, SIGNAL = 1.2
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 1122
CQ108842
LOCUS       CQ108842                152 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17701 from Patent WO0157272.
ACCESSION   CQ108842
VERSION     CQ108842.1  GI:41077895
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 17701 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN PLACENTA, SIGNAL = 7.1
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||| |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 1123
CQ147474
LOCUS       CQ147474                152 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17496 from Patent WO0157276.
ACCESSION   CQ147474
VERSION     CQ147474.1  GI:41154824
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 17496 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.2
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
```

```
              ||||||||:|||||| ||||:|||||||||| ||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTAGCGGAA 143


RESULT 1124
CQ207262
LOCUS       CQ207262                152 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 9539 from Patent WO0157271.
ACCESSION   CQ207262
VERSION     CQ207262.1  GI:41193431
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 9539 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN BT474, SIGNAL = 0.95
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTAGCGGAA 143


RESULT 1125
CQ230720
LOCUS       CQ230720                152 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 17559 from Patent WO0157273.
ACCESSION   CQ230720
VERSION     CQ230720.1  GI:41213938
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 17559 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTAGCGGAA 143


RESULT 1126
CQ268859
LOCUS       CQ268859                152 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 17120 from Patent WO0157277.
ACCESSION   CQ268859
VERSION     CQ268859.1  GI:41241463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 17120 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              |||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
```

```
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 1127
CQ305886
LOCUS       CQ305886                152 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 16991 from Patent WO0186003.
ACCESSION   CQ305886
VERSION     CQ305886.1  GI:41266463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 16991 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .152
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL022067.1
                     EXPRESSED IN LUNG, SIGNAL = 1.3
                     NT HIT: gi4759007, EVALUE 3.00e-54
                     EST_HUMAN HIT: AU138097.1, EVALUE 4.00e-54
                     SWISSPROT HIT: P10269, EVALUE 5.00e-10"

  Query Match             1.6%;  Score 122.6;  DB 1;  Length 152;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  128;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       6674 TGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRAAATCGAAGAG 6733
              ||||||||||||||||||||||||||| ||||||| ||||||||||||:|||| ||||||
Db          1 TGGTCACCTTGCAAGATCAACTTAACTTCCTAGCATCAGTAGTCCTTCAAAATTGAAGAG 60

Qy       6734 CTTTAGACTYGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTTAGGGGAAGAATGCT 6793
              |||||||||:||||||| ||||:|||||||||| |||||||||||||||||||||||| |
Db         61 CTTTAGACTTGCTAACCTCTGAAAGAGGGGGAAGCTGTTTATTTTTAGGGGAAGAATGTT 120

Qy       6794 GTTATTATGTTAATCAATCCGGA 6816
              ||||||||||||  |  | | | |
Db        121 GTTATTATGTTATTTTAGCGGAA 143


RESULT 1128
CQ117238
LOCUS       CQ117238                133 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 26097 from Patent WO0157272.
ACCESSION   CQ117238
VERSION     CQ117238.1  GI:41087108
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 26097 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .133
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007567.1
                     EXPRESSED IN PLACENTA, SIGNAL = 2.7
                     SWISSPROT HIT: P21414, EVALUE 1.00e-05
                     EST_HUMAN HIT: AU121943.1, EVALUE 2.00e-52
                     NT HIT: AF009668.1, EVALUE 2.00e-57"

  Query Match             1.6%;  Score 121.8;  DB 1;  Length 133;
  Best Local Similarity   90.2%;  Pred. No. 0;
  Matches  120;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2857 ATACTTCAAACCCARCATCTCAACTCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATA 2916
              || |||| ||||||:| |||||||||| ||||||:|||||||||||||||||||||||||
Db          1 ATCCTTCGAACCCAACGTCTCAACTCACTTGGACTGTTTTACCCCAAGGGTTCAGGGATA 60

Qy       2917 GYCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGYCARTYMTCATACCTGGACA 2976
              |:|||||||||||||||| ||||||||||||||||||:||:|::||||||||||||||
Db         61 GCCCCCATCTATTTGGCCAGACATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGACA 120

Qy       2977 CTCTTGTCCTTCR 2989
              ||||||||||||:
Db        121 CTCTTGTCCTTCA 133


RESULT 1129
CQ745278/c
LOCUS       CQ745278                168 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31212 from Patent WO02068579.
ACCESSION   CQ745278
VERSION     CQ745278.1  GI:42363155
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 31212 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .168
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.6%;  Score 121.4;  DB 1;  Length 168;
  Best Local Similarity   82.0%;  Pred. No. 0;
  Matches  137;  Conservative    1;  Mismatches   29;  Indels    0;  Gaps    0;

Qy        147 TCGAGCTGAGCTTTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCG 206
              || | |||||||||  ||| || |||| |||||||||||||| |   | || | | |||
Db        168 TCAAACTGAGCTTTCACTCGCCATCCAACACTGCTGTTTGCCGCTGTCACAGATCCGCCA 109

Qy        207 CTGACTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCG 266
              |||| | ||||||||| |||||| |||| |||||||||||||||||| ||||||:| |
Db        108 CTGATTTCCATCCCTCAGGATCCGGCAGAGTGTCCGCTGTGCTCCTGATTCAGCGAGGAG 49

Qy        267 CCCATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              ||||||||| ||||||||  |||||||||||||  ||||| ||||||
Db         48 CCCATTGCCACTCCCAATCAGGCTAAAGGCTTGTAATTGTTCCTGCA 2


RESULT 1130
CQ075479
LOCUS       CQ075479                155 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 11279 from Patent WO0157278.
ACCESSION   CQ075479
```

```
VERSION     CQ075479.1  GI:41045348
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 11279 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN HELA, SIGNAL = 2.2
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1131
CQ106452
LOCUS       CQ106452                155 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 15311 from Patent WO0157272.
ACCESSION   CQ106452
VERSION     CQ106452.1  GI:41075505
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 15311 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN PLACENTA, SIGNAL = 8.6
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1132
CQ145113
LOCUS       CQ145113                155 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 15135 from Patent WO0157276.
ACCESSION   CQ145113
VERSION     CQ145113.1  GI:41102485
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 15135 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.2
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | ||||||||| ||||||| ||||||| || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1133
CQ180556
LOCUS       CQ180556                155 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 11952 from Patent WO0157274.
ACCESSION   CQ180556
VERSION     CQ180556.1  GI:41175295
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 11952 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN HEART, SIGNAL = 1.3
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
              ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              || | ||||||||| |||||||  |||||||  || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
              | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1134
CQ228295
LOCUS       CQ228295                155 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 15134 from Patent WO0157273.
ACCESSION   CQ228295
VERSION     CQ228295.1  GI:41211513
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 15134 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
              ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              || | ||||||||| |||||||  |||||||  || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
              | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1135
CQ303441
LOCUS       CQ303441                155 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 14546 from Patent WO0186003.
ACCESSION   CQ303441
VERSION     CQ303441.1  GI:41264018
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 14546 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN LUNG, SIGNAL = 2.1
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
              ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:|||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
              || | ||||||||| |||||||  |||||||  || | |||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
              | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1136
CQ340739
LOCUS       CQ340739                155 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 14833 from Patent WO0157275.
ACCESSION   CQ340739
VERSION     CQ340739.1  GI:41289810
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 14833 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL121572.4
                     EXPRESSED IN BRAIN, SIGNAL = 0.92
                     EST_HUMAN HIT: AU121481.1, EVALUE 1.00e-32
                     NT HIT: AB026898.1, EVALUE 8.00e-42
                     SWISSPROT HIT: P52490, EVALUE 6.60e+00"

  Query Match             1.6%;  Score 121;  DB 1;  Length 155;
  Best Local Similarity   85.8%;  Pred. No. 0;
  Matches  133;  Conservative    1;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       7043 AGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTGYTAGCC 7102
             ||||| |||||||||||| | |||||||||||| | ||||||||||||||||||:||||||
Db          1 AGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTGCTAGCC 60

Qy       7103 CACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCACAAC 7162
             || | |||||||| |||||||  |||||||  || | ||||||||||| |||||| ||
Db         61 CATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCACGAC 120

Qy       7163 CTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAG 7197
             | ||||||| |||||||| ||| ||||||| ||||
Db        121 CCCTACTACACCCCAATTTAGCGGGAAGCAATTAG 155


RESULT 1137
CQ748116/c
LOCUS       CQ748116                171 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34050 from Patent WO02068579.
ACCESSION   CQ748116
VERSION     CQ748116.1  GI:42372931
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 34050 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .171
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.6%;  Score 120.6;  DB 1;  Length 171;
  Best Local Similarity   85.2%;  Pred. No. 0;
  Matches  132;  Conservative    1;  Mismatches   22;  Indels    0;  Gaps    0;

Qy        159 TTTGCTCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATC 218
              ||| ||| || |||||||||||||||| || ||  |||  ||| |||||||||| |||||
Db        156 TTTCCTCGCCATCCACCACTGCTGTTTTCCGCCGTCGCGGACCCGCCGCTGACTTCCATC 97

Qy        219 CCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCATTGCCGCT 278
              ||||  ||||| |||||||||| |||||||||||||||||| |:|||||||||| |||
Db         96 CCTCCAGATCCGGCAGGGTGTCCACTGTGCTCCTGATCCAGCAAGGCGCCCATTGCTGCT 37

Qy        279 CCCAATTGGGCTAAAGGCTTGCCATTGTNCCTGCA 313
              ||  || |||||||||||| |||||||| ||||||
Db         36 CCTGATGGGGCTAAAGGCTCGCCATTGTTCCTGCA 2


RESULT 1138
CQ742851
LOCUS       CQ742851                149 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28785 from Patent WO02068579.
ACCESSION   CQ742851
VERSION     CQ742851.1  GI:42354174
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 28785 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .149
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.6%;  Score 119.3;  DB 1;  Length 149;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  131;  Conservative    7;  Mismatches    9;  Indels    1;  Gaps    1;

Qy       3801 CAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCG 3860
             ||||||||||||||||:||:|||||||||||||| :| ||| |||||||||||||  |||
Db          2 CAATGAAGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCTGCTCA 61

Qy       3861 AGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA 3920
             | ||||||||:||||:||||||:||||||||||||:|||| ||||||||||||||||||||
Db         62 AAGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAA 120

Qy       3921 GTTCCTTTGTAGAAAAAGGACTTCGAAA 3948
             ||||||| |||||||||||||||| |||
Db        121 GTTCCTTCGTAGAAAAAGGACTTCAAAA 148


RESULT 1139
CQ748132
LOCUS       CQ748132                149 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34066 from Patent WO02068579.
ACCESSION   CQ748132
VERSION     CQ748132.1  GI:42372987
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 34066 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .149
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.6%;  Score 119.3;  DB 1;  Length 149;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches  131;  Conservative    7;  Mismatches    9;  Indels    1;  Gaps    1;
```

```
Qy       3801 CAATGAAGAAAAGATARAAYATAACTGTCAACAARTAATTTCTCAAACCTATGCCACTCG 3860
              ||||||||||||||||:||:|||||||||||||| :| ||| ||||||||||||  |||
Db          2 CAATGAAGAAAAGATAGAACATAACTGTCAACAGGTGATTGCTCAAACCTATGCTGCTCA 61

Qy       3861 AGGGGACCTTYTAGARGTTCCYTTGACTGATCCYGACCTTCAACTTGTATACTGATGGAA 3920
              | ||||||||:||||:|||||:||||||||||:|||| ||||||||||||||||||||||
Db         62 AAGGGACCTTCTAGAGGTTCCCTTGACTGATCCCGACC-TCAACTTGTATACTGATGGAA 120

Qy       3921 GTTCCTTTGTAGAAAAAGGACTTCGAAA 3948
              ||||||| |||||||||||||||| |||
Db        121 GTTCCTTCGTAGAAAAAGGACTTCAAAA 148


RESULT 1140
AF009667
LOCUS       AF009667                 118 bp    mRNA    linear   VRL 12-AUG-1997
DEFINITION  Multiple sclerosis associated retrovirus conserved pol gene
            sequence.
ACCESSION   AF009667
VERSION     AF009667.1  GI:2323285
KEYWORDS    .
SOURCE      Multiple sclerosis associated retrovirus
  ORGANISM  Multiple sclerosis associated retrovirus
            Viruses; Retro-transcribing viruses; Retroviridae.
REFERENCE   1  (bases 1 to 118)
  AUTHORS   Tuke,P.W., Perron,H., Bedin,F., Beseme,F., Mallet,F. and
            Garson,J.A.
  TITLE     Development of a pan-retrovirus detection system for multiple
            sclerosis studies
  JOURNAL   Acta Neurol. Scand., Suppl. S169, 16-21 (1997)
REFERENCE   2  (bases 1 to 118)
  AUTHORS   Perron,H., Garson,J.A., Bedin,F., Beseme,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Mallet,F., Tuke,P.W., Voisset,C., Blond,J.L.,
            Lalande,B., Seigneurin,J.M., Mandrand,B. and The collaborative
            research group on MS.
  TITLE     Molecular identification of a novel retrovirus repeatedly isolated
            from patients with multiple sclerosis. The Collaborative Research
            Group on Multiple Sclerosis
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 94 (14), 7583-7588 (1997)
   PUBMED   9207135
REFERENCE   3  (bases 1 to 118)
  AUTHORS   Perron,H., Bedin,F., Beseme,F., Mallet,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Garson,J.A. and Tuke,P.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1997) UMR103 / E.N.S., BIOMERIEUX, 46, allee
            d'Italie, Lyon CEDEX 07 69364, France
COMMENT     International patent WO-97/06260 and other patents pending.
FEATURES             Location/Qualifiers
     source          1. .118
                     /organism="Multiple sclerosis associated retrovirus"
                     /mol_type="mRNA"
                     /db_xref="taxon:62805"
     misc_feature    1. .118
                     /note="conserved pol sequence"

  Query Match             1.5%;  Score 110.6;  DB 1;  Length 118;
  Best Local Similarity   91.5%;  Pred. No. 0;
  Matches  107;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
              :| || |||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db          2 GTCTTGCCCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 61

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTT 3008
              |||||:||:|::|||||||||||||||||||||||:|||:|||||||||||||||
Db         62 TTGAGTCAATTCTCATACCTGGACACTCTTGTCCTTCAGTATGTGGATGATTTACTT 118


RESULT 1141
HUMSWX525/c
LOCUS       HUMSWX525                139 bp    DNA     linear   STS 13-OCT-1993
DEFINITION  Human chromosome X STS sWXD525, single read, sequence tagged site.
ACCESSION   L24603
VERSION     L24603.1  GI:405353
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 139)
  AUTHORS   Schlessinger,D.
  TITLE     Human chromosome X STS
  JOURNAL   Unpublished (1993)
COMMENT     Original source text: Homo sapiens DNA.
            Submitted by: David Schlessinger,
            Center for Genetics in Medicine,
            Washington University School of Medicine, Box 8232 4566 Scott
            Avenue, St. Louis, MO 63110, USA
            e-mail: davids@wugenmail.wustl.edu
            Primer A:    CATGCACTGAGTGTTATGG
            Primer B:    CGGGTTCATCCTAATCAAG
            STS Size:    97 bp
            Template:        250 ng/100ul
            Primer:          45 pmoles/100ul
            dNTPs:           100 uM
            MgCl2:           1.5 mM
            KCl:             100 mM
            Tris-HCl:        10 mM
            Taq Polymerase:  0.1 U
            NH4Cl:           5 mM
            pH:              8.6
            Total Vol:       10 ul
            PCR Profile:
            Denaturation:     94 degrees C for 0.50 minute(s)
            Annealing:        55 degrees C for 0.75 minute(s)
            Polymerization:   72 degrees C for 0.75 minute(s)
            PCR Cycles:       30
            Thermal Cycler:   PE 9600.
FEATURES             Location/Qualifiers
     source          1. .139
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="Xq25-q26"
     STS             21. .117
                     /standard_name="sWXD525"
     primer_bind     21. .39
     primer_bind     complement(99. .117)

  Query Match             1.5%;  Score 109.4;  DB 1;  Length 139;
  Best Local Similarity   85.6%;  Pred. No. 0;
  Matches  119;  Conservative    1;  Mismatches   19;  Indels    0;  Gaps    0;

Qy        304 TGTNCCTGCACGGCTAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACTGGG 363
              ||| |||||||||||||||||| |||||  | |||||  |||||||||||| ||||||||
Db        139 TGTTCCTGCACGGCTAAGTGCCCGGGTTCATCCTAATCAAGCTGAACACTAGTCACTGGG 80

Qy        364 TTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGG 423
              ||||| |||||||||| |||||||| ||||||||||:| | ||||||| |   |||||
Db         79 TTCCACGGTTCTCTTCCGTGACCCATGGCTTCTAATAGAGCCATAACACTCAGTGCATGG 20

Qy        424 CCCAAGATTCCATTCCTTG 442
              |||| ||||||||| ||||
Db         19 CCCAGGATTCCATTTCTTG 1


RESULT 1142
CQ082276/c
LOCUS       CQ082276                 135 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 18076 from Patent WO0157278.
ACCESSION   CQ082276
VERSION     CQ082276.1  GI:41052145
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 18076 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN HELA, SIGNAL = 3.9
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| ||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 1143
CQ117002/c
LOCUS       CQ117002                  135 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25861 from Patent WO0157272.
ACCESSION   CQ117002
VERSION     CQ117002.1  GI:41086872
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 25861 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN PLACENTA, SIGNAL = 2.5
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| ||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 1144
CQ155677/c
LOCUS       CQ155677                  135 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25699 from Patent WO0157276.
ACCESSION   CQ155677
VERSION     CQ155677.1  GI:41163029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 25699 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN BONE MARROW, SIGNAL = 3.2
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
              :||||||| | ||||||||| ||||||| ||||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
              ||  ||| |||||| |||||| |||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
              | | |
Db         15 CTCAA 11


RESULT 1145
CQ188073/c
LOCUS       CQ188073                  135 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 19469 from Patent WO0157274.
ACCESSION   CQ188073
VERSION     CQ188073.1  GI:41183085
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
```

```
  JOURNAL   Patent: WO 0157274-A 19469 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN HEART, SIGNAL = 2
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 1146
CQ238853/c
LOCUS       CQ238853                 135 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 25692 from Patent WO0157273.
ACCESSION   CQ238853
VERSION     CQ238853.1  GI:41222244
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 25692 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 1147
CQ276533/c
LOCUS       CQ276533                 135 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24794 from Patent WO0157277.
ACCESSION   CQ276533
VERSION     CQ276533.1  GI:41249137
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 24794 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | ||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 1148
CQ313549/c
LOCUS       CQ313549                 135 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 24654 from Patent WO0186003.
ACCESSION   CQ313549
VERSION     CQ313549.1  GI:41274126
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 24654 15-NOV-2001;
```

```
           Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN LUNG, SIGNAL = 1.9
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 1149
CQ351031/c
LOCUS       CQ351031                135 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 25125 from Patent WO0157275.
ACCESSION   CQ351031
VERSION     CQ351031.1  GI:41300102
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 25125 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL139245.2
                     EXPRESSED IN BRAIN, SIGNAL = 2.2
                     SWISSPROT HIT: P18080, EVALUE 2.40e+00
                     NT HIT: gi7656876, EVALUE 2.00e-33
                     EST_HUMAN HIT: AU158595.1, EVALUE 1.00e-35"

  Query Match             1.4%;  Score 101.8;  DB 1;  Length 135;
  Best Local Similarity   87.2%;  Pred. No. 0;
  Matches  109;  Conservative    2;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       7097 YTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTG 7156
             :||||||| | |||||||||| |||||||  |||||||||||||||||||||||||| |||
Db        135 CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACCCCTCCTGAGGAAATCTCAACTG 76

Qy       7157 CACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCCAACCT 7216
             ||  ||| |||||| |||||| ||||||||||||||||||||||||: ||||||||||
Db         75 CATGACCCCTACTATGCCCCAGTTCAGCAGGAAGCAGTTAGAGCGGTCATCGGCCAACCT 16

Qy       7217 CCCCA 7221
             | | |
Db         15 CTCAA 11


RESULT 1150
CQ751084/c
LOCUS       CQ751084                113 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 37018 from Patent WO02068579.
ACCESSION   CQ751084
VERSION     CQ751084.1  GI:42383553
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 37018 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .113
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.3%;  Score 100.4;  DB 1;  Length 113;
  Best Local Similarity   92.0%;  Pred. No. 0;
  Matches  104;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        336 CTAATTGAGCTGAACACTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTC 395
              ||||| |||||||||||||| |||||||||||||||||||||||||  |||||||| |||||
Db        113 CTAATCGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCCATGACCCACAGCTTC 54

Qy        396 TAATAKAACTATAACACTTACCACATGGCCCAAGATTCCATTCCTTGGAATCC 448
              ||| |:| |||||||||| ||||||||||||||||||||||||||||||||||
Db         53 TAACAGAGCTATAACACTCACCACATGGCCCAAGATTCCATTCCTTGGAATCC 1


RESULT 1151
CQ113073/c
LOCUS       CQ113073                110 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 21932 from Patent WO0157272.
ACCESSION   CQ113073
VERSION     CQ113073.1  GI:41082943
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 21932 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN PLACENTA, SIGNAL = 4.6
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             ||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1152
CQ151943/c
LOCUS       CQ151943                110 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 21965 from Patent WO0157276.
ACCESSION   CQ151943
VERSION     CQ151943.1  GI:41159293
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 21965 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN BONE MARROW, SIGNAL = 2.5
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1153
CQ185508/c
LOCUS       CQ185508                110 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 16904 from Patent WO0157274.
ACCESSION   CQ185508
VERSION     CQ185508.1  GI:41180523
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 16904 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN HEART, SIGNAL = 2.2
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1154
CQ235300/c
LOCUS       CQ235300                110 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 22139 from Patent WO0157273.
ACCESSION   CQ235300
VERSION     CQ235300.1  GI:41218577
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 22139 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
             |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1155
CQ272864/c
LOCUS       CQ272864                110 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 21125 from Patent WO0157277.
```

```
ACCESSION   CQ272864
VERSION     CQ272864.1  GI:41245468
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 21125 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1156
CQ310432/c
LOCUS       CQ310432                110 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 21537 from Patent WO0186003.
ACCESSION   CQ310432
VERSION     CQ310432.1  GI:41271009
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 21537 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN LUNG, SIGNAL = 2.4
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1157
CQ347148/c
LOCUS       CQ347148                110 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 21242 from Patent WO0157275.
ACCESSION   CQ347148
VERSION     CQ347148.1  GI:41296219
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 21242 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL049870.3
                     EXPRESSED IN BRAIN, SIGNAL = 2.3
                     NT HIT: AF208161.1, EVALUE 6.00e-36
                     SWISSPROT HIT: P03384, EVALUE 7.00e-04
                     EST_HUMAN HIT: T70359.1, EVALUE 9.00e-36"

  Query Match             1.2%;  Score 93;  DB 1;  Length 110;
  Best Local Similarity   91.4%;  Pred. No. 0;
  Matches   96;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6938 GACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTRACCTCCTTG 6997
             |||||||||||||||||||||||| ||||| |||||||||||||||||||:||||||||
Db        110 GACCTCTAGCAGCTATAATATTGTTACTCTTCTTTGGACCCTGTATCTTTAACCTCCTTA 51

Qy       6998 TTAACTTTGTCTCTTCCAGAATCGAAGCTGTRAAACTACAAATGG 7042
              |||| || |||||||||||||||||||| ||:|| ||||||||||
Db         50 TTAAATTGGTCTCTTCCAGAATCGAAGCGGTAAAGCTACAAATGG 6


RESULT 1158
A57086
LOCUS       A57086                  126 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 39 from Patent WO9628552.
ACCESSION   A57086
VERSION     A57086.1  GI:3713069
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 39 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .126
                     /organism="unidentified"
```

```
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.2%;  Score 92;  DB 1;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0;
  Matches   98;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
                :|  |||| || |||||||||||||||: |||||||||||||||| | ||||| ||| ||||
Db          1 GTGCTACCACAGGGGTTCAGGGATAGCTCCCATCTATTTGGCCTGACATTAACCCGAGAC 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTR 3011
                || ||:||:|::|||||| |||||||||||||||| :||| |||||||  | ||  |:
Db         61 TTAAGCCAGTTCTCATACGTGGACACTCTTGTCCTTTGGTACGTGGATGACATCCTGCTG 120

Qy       3012 GCYGCC 3017
                ||: ||
Db        121 GCCTCC 126


RESULT 1159
A80344
LOCUS       A80344                 126 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 39 from Patent EP0731168.
ACCESSION   A80344
VERSION     A80344.1  GI:6093071
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 126)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 39 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .126
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.2%;  Score 92;  DB 1;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0;
  Matches   98;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
                :|  |||| || |||||||||||||||: |||||||||||||||| | ||||| ||| ||||
Db          1 GTGCTACCACAGGGGTTCAGGGATAGCTCCCATCTATTTGGCCTGACATTAACCCGAGAC 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGATTTACTTTTR 3011
                || ||:||:|::|||||| |||||||||||||||| :||| |||||||  | ||  |:
Db         61 TTAAGCCAGTTCTCATACGTGGACACTCTTGTCCTTTGGTACGTGGATGACATCCTGCTG 120

Qy       3012 GCYGCC 3017
                ||: ||
Db        121 GCCTCC 126


RESULT 1160
CQ057462/c
LOCUS       CQ057462               103 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 8282 from Patent WO0157270.
ACCESSION   CQ057462
VERSION     CQ057462.1  GI:41031968
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 8282 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN HBL100, SIGNAL = 1.4
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                     NT HIT: AF009668.1, EVALUE 1.00e-36
                     SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
                ||  |||||||||||||||||||||| | |||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
                ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1161
CQ076748/c
LOCUS       CQ076748               103 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 12548 from Patent WO0157278.
ACCESSION   CQ076748
VERSION     CQ076748.1  GI:41046617
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 12548 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN HELA, SIGNAL = 1.6
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                     NT HIT: AF009668.1, EVALUE 1.00e-36
                     SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
                ||  |||||||||||||||||||||| | |||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
                ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1162
CQ107735/c
```

```
LOCUS       CQ107735                103 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 16594 from Patent WO0157272.
ACCESSION   CQ107735
VERSION     CQ107735.1  GI:41076788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 16594 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN PLACENTA, SIGNAL = 3.3
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                     NT HIT: AF009668.1, EVALUE 1.00e-36
                     SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||||  |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||| |||||||| ||||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1163
CQ146384/c
LOCUS       CQ146384                103 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 16406 from Patent WO0157276.
ACCESSION   CQ146384
VERSION     CQ146384.1  GI:41153734
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 16406 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.3
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                     NT HIT: AF009668.1, EVALUE 1.00e-36
                     SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||||  |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||| |||||||| ||||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1164
CQ206196/c
LOCUS       CQ206196                103 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 8473 from Patent WO0157271.
ACCESSION   CQ206196
VERSION     CQ206196.1  GI:41192300
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 8473 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AC007244.1
                     EXPRESSED IN BT474, SIGNAL = 1.3
                     EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                     NT HIT: AF009668.1, EVALUE 1.00e-36
                     SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
             ||  |||||||||||||||||||||| | |||||||||||||  |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
             |||||||| |||||||| ||||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1165
CQ229597/c
LOCUS       CQ229597                103 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 16436 from Patent WO0157273.
ACCESSION   CQ229597
VERSION     CQ229597.1  GI:41212815
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
```

```
          Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 16436 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .103
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN ADULT LIVER, SIGNAL = 2
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                      NT HIT: AF009668.1, EVALUE 1.00e-36
                      SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
              ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1166
CQ267736/c
LOCUS       CQ267736                 103 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 15997 from Patent WO0157277.
ACCESSION   CQ267736
VERSION     CQ267736.1  GI:41240340
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 15997 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .103
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                      SWISSPROT HIT: P10272, EVALUE 8.00e-05
                      NT HIT: AF009668.1, EVALUE 1.00e-36"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
              ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1167
CQ304760/c
LOCUS       CQ304760                 103 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 15865 from Patent WO0186003.
ACCESSION   CQ304760
VERSION     CQ304760.1  GI:41265337
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 15865 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .103
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN LUNG, SIGNAL = 1.2
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                      NT HIT: AF009668.1, EVALUE 1.00e-36
                      SWISSPROT HIT: P10272, EVALUE 8.00e-05"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
              ||||||||| ||||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1168
CQ342005/c
LOCUS       CQ342005                 103 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 16099 from Patent WO0157275.
ACCESSION   CQ342005
VERSION     CQ342005.1  GI:41291076
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 16099 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .103
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AC007244.1
                      EXPRESSED IN BRAIN, SIGNAL = 1.2
                      EST_HUMAN HIT: AA368311.1, EVALUE 3.00e-34
                      SWISSPROT HIT: P10272, EVALUE 8.00e-05
                      NT HIT: AF009668.1, EVALUE 1.00e-36"

  Query Match             1.2%;  Score 89.2;  DB 1;  Length 103;
  Best Local Similarity   92.2%;  Pred. No. 0;
  Matches   94;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2275 TTCTTTTCACATGCTTTTCTAATTATGCTTGAAAGCCCCACTACCTTGTTAGGGAGAGAC 2334
              ||  |||||||||||||||||||| | ||||||||||||| |||| ||||||||||||
```

```
Db        102 TTACTTTCACATGCTTTTCTAATTATACCTGAAAGCCCCACTCCCTTATTAGGGAGAGAC 43

Qy       2335 ATTCTAGCAAAAGCAGGGGCCATTATACACCTGAACATAGGA 2376
              ||||||||| |||||||| ||||||||||||||||||||||
Db         42 ATTCTAGCAGAAGCAGGGGTCATTATACACCTGAACATAGGA 1


RESULT 1169
CQ055176
LOCUS       CQ055176                107 bp    DNA     linear   PAT 19-JAN-2004
DEFINITION  Sequence 5996 from Patent WO0157270.
ACCESSION   CQ055176
VERSION     CQ055176.1  GI:41029682
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and hbl 100 cells
  JOURNAL   Patent: WO 0157270-A 5996 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HBL100, SIGNAL = 1.2
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
              || |||  |||||||||||||| |||||||  | ||||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1170
CQ074435
LOCUS       CQ074435                107 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 10235 from Patent WO0157278.
ACCESSION   CQ074435
VERSION     CQ074435.1  GI:41044304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 10235 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HELA, SIGNAL = 1.3
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
              || |||  |||||||||||||| |||||||  | ||||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1171
CQ105333
LOCUS       CQ105333                107 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14192 from Patent WO0157272.
ACCESSION   CQ105333
VERSION     CQ105333.1  GI:41074386
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
  JOURNAL   Patent: WO 0157272-A 14192 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN PLACENTA, SIGNAL = 5
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
              || |||  |||||||||||||| |||||||  | ||||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
              ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1172
CQ144027
LOCUS       CQ144027                107 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14049 from Patent WO0157276.
ACCESSION   CQ144027
VERSION     CQ144027.1  GI:41101399
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human bone marrow
  JOURNAL   Patent: WO 0157276-A 14049 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN BONE MARROW, SIGNAL = 1.6
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1173
CQ179507
LOCUS       CQ179507                107 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 10903 from Patent WO0157274.
ACCESSION   CQ179507
VERSION     CQ179507.1  GI:41174246
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
  JOURNAL   Patent: WO 0157274-A 10903 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN HEART, SIGNAL = 1.4
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1174
CQ203866
LOCUS       CQ203866                107 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 6143 from Patent WO0157271.
ACCESSION   CQ203866
VERSION     CQ203866.1  GI:41189713
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human breast and bt 474 cells
  JOURNAL   Patent: WO 0157271-A 6143 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN BT474, SIGNAL = 1.6
                     NT HIT: AB026898.1, EVALUE 1.00e-27
                     EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                     SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
                || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
                ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1175
CQ227234
LOCUS       CQ227234                107 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 14073 from Patent WO0157273.
ACCESSION   CQ227234
VERSION     CQ227234.1  GI:41210437
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
            August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
            (03.10.00)<150> US 60/236,359<151> 27 September 2000
            (27.09.00)<150> US 60/234,687<151> 21 September 2000
            (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
            Molecular Dynamics Sequence Listing Engine
  JOURNAL   Patent: WO 0157273-A 14073 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL008629.9
                     EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
```

```
                  NT HIT: AB026898.1, EVALUE 1.00e-27
                  EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                  SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1176
CQ265383
LOCUS       CQ265383                 107 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13644 from Patent WO0157277.
ACCESSION   CQ265383
VERSION     CQ265383.1  GI:41237971
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human fetal liver
  JOURNAL   Patent: WO 0157277-A 13644 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .107
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL008629.9
                      EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
                      NT HIT: AB026898.1, EVALUE 1.00e-27
                      EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                      SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1177
CQ302469
LOCUS       CQ302469                 107 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13574 from Patent WO0186003.
ACCESSION   CQ302469
VERSION     CQ302469.1  GI:41263046
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human lung
  JOURNAL   Patent: WO 0186003-A 13574 15-NOV-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .107
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL008629.9
                      EXPRESSED IN LUNG, SIGNAL = 1.5
                      NT HIT: AB026898.1, EVALUE 1.00e-27
                      EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                      SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1178
CQ339647
LOCUS       CQ339647                 107 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 13741 from Patent WO0157275.
ACCESSION   CQ339647
VERSION     CQ339647.1  GI:41288718
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human brain
  JOURNAL   Patent: WO 0157275-A 13741 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .107
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
                      /note="MAP TO AL008629.9
                      EXPRESSED IN BRAIN, SIGNAL = 1.3
                      NT HIT: AB026898.1, EVALUE 1.00e-27
                      EST_HUMAN HIT: AA368311.1, EVALUE 5.00e-18
                      SWISSPROT HIT: O35083, EVALUE 2.40e+00"

  Query Match             1.1%;  Score 85.2;  DB 1;  Length 107;
  Best Local Similarity   87.7%;  Pred. No. 0;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2471 CCGTCCTGTTCAAGTTAAACTAAAGGATTCCACTTCCTTTCCCTACCAAAGGCAGTACCC 2530
             || |||  |||||||||||||| |||||||  | |||||||||||||||||||||||| |
Db          1 CCATCCCATTCAAGTTAAACTACAGGATTCTGCCTCCTTTCCCTACCAAAGGCAGTACTC 60

Qy       2531 CCTCAGACCCAAGGCCCAACAAGGATTCCAAAAGATTGTTAAGGAC 2576
             ||| |||||| ||||| | |||||| ||||||||||||||||||||
Db         61 CCTTAGACCCGAGGCCGATCAAGGACTCCAAAAGATTGTTAAGGAC 106


RESULT 1179
CQ751265/c
LOCUS       CQ751265                 108 bp    DNA     linear   PAT 03-FEB-2004
```

```
DEFINITION  Sequence 37199 from Patent WO02068579.
ACCESSION   CQ751265
VERSION     CQ751265.1  GI:42412021
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 37199 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

  Query Match             1.1%;  Score 82.6;  DB 1;  Length 108;
  Best Local Similarity   84.3%;  Pred. No. 0;
  Matches   91;  Conservative    1;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        164 TCACCGTCCACCACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCT 223
              | |||||||||||| ||||||||| ||  | || |||||| ||||||| ||| |||||
Db        108 TTACCGTCCACCACAGCTGTTTGCCGCCGTCACAGACCTGCTGCTGACTTCCACCCCTCC 49

Qy        224 GGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGARGCGCCCAT 271
                  ||||  ||||||||||||||||||||||||||||||:||  ||||
Db         48 AGATCTGGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCATCCAT 1


RESULT 1180
HS97F2F
LOCUS       HS97F2F                 100 bp    DNA     linear   PRI 22-OCT-1995
DEFINITION  H.sapiens CpG island DNA genomic Mse1 fragment, clone 97f2, forward
            read cpg97f2.ft1a.
ACCESSION   Z64129
VERSION     Z64129.1  GI:1036507
KEYWORDS    CpG island; genomic Mse1 fragment.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE     Purification of CpG islands using a methylated DNA binding column
  JOURNAL   Nat. Genet. 6 (3), 236-244 (1994)
    PUBMED  8012384
REFERENCE   2  (bases 1 to 100)
  AUTHORS   MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT     Vector: pGEM-5Zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="97f2"
                     /sex="male"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /dev_stage="adult"

  Query Match             1.1%;  Score 81.1;  DB 1;  Length 100;
  Best Local Similarity   90.1%;  Pred. No. 0;
  Matches   91;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       3615 TAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACT 3674
              ||||| |||||| ||:|||||||||||||||||||| |||||:|||||::|||||||
Db          1 TAAAAGAATACAAGGAAGAGATCTTACTGTGTGGACATC-CATGATGTGAATGGCATACT 59

Qy       3675 CACTGCTAAAGGAGACTTGTGGCTGTCAGACAACYGTTTAC 3715
              |||||||||||||||||||||||||||||||||:  ||||
Db         60 CACTGCTAAAGGAGACTTGTGGCTGTCAGACAACCACTTAC 100


RESULT 1181
A46018
LOCUS       A46018                  111 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 7 from Patent EP0674004.
ACCESSION   A46018
VERSION     A46018.1  GI:2300290
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Mandrand,B., Perron,H., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and pathogenic and/or infections agent MSRV2 linked to
            multiple sclerosis, their nucleic components and their applications
  JOURNAL   Patent: EP 0674004-A 7 27-SEP-1995;
            BIO MERIEUX (FR)
COMMENT     Other publication DE 674004T 960919
            Other publication NO 953925 951204
            Other publication FI 954699 951003
            Other publication AU 1711495 950821
            Other publication CA 2141907 950805
            Other publication FR 2715939 950811
            Other publication FR 2715938 950811
            Other publication FR 2715937 950811
            Other publication FR 2715936 950811.
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
              :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
              :|::||||:|::|:|:::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1182
A57054
LOCUS       A57054                  111 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 7 from Patent WO9628552.
ACCESSION   A57054
VERSION     A57054.1  GI:3713037
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 7 19-SEP-1996;
```

```
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES              Location/Qualifiers
     source           1. .111
                      /organism="unidentified"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1183
A60119
LOCUS       A60119                  111 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 7 from Patent WO9706260.
ACCESSION   A60119
VERSION     A60119.1  GI:3715135
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC
            PURPOSES
  JOURNAL   Patent: WO 9706260-A 7 20-FEB-1997;
            BIO MERIEUX (FR)
COMMENT     Other publication FR 2737500 970207.
FEATURES              Location/Qualifiers
     source           1. .111
                      /organism="unidentified"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1184
A79469
LOCUS       A79469                  111 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 7 from Patent EP0789077.
ACCESSION   A79469
VERSION     A79469.1  GI:6092477
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Jolivet-Reynaud,C. and Bedin,F.
  TITLE     VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE
            SCLEROSE HAVING DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTICAL USES
  JOURNAL   Patent: EP 0789077-A 7 13-AUG-1997;
            BIO MERIEUX (FR)
FEATURES              Location/Qualifiers
     source           1. .111
                      /organism="unidentified"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1185
A80203
LOCUS       A80203                  111 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 7 from Patent WO9521256.
ACCESSION   A80203
VERSION     A80203.1  GI:6092935
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H. and Mallet,F.
  TITLE     MSRV1 VIRUS AND MSRV2 PATHOGENIC AND/OR INFECTIOUS AGENT ASSOCIATED
            WITH MULTIPLE SCLEROSIS, NUCLEIC ACID COMPONENTS AND APPLICATIONS
            OF SAME
  JOURNAL   Patent: WO 9521256-A 7 10-AUG-1995;
            BIO MERIEUX (FR); PERRON HERVE (FR)
FEATURES              Location/Qualifiers
     source           1. .111
                      /organism="unidentified"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1186
A80312
LOCUS       A80312                  111 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 7 from Patent EP0731168.
ACCESSION   A80312
VERSION     A80312.1  GI:6093039
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H. and Bedin,F.
```

```
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 7 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1187
AR035355
LOCUS       AR035355                111 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5871745.
ACCESSION   AR035355
VERSION     AR035355.1  GI:5952023
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Multiple sclerosis related virus
  JOURNAL   Patent: US 5871745-A 7 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1188
AR036196
LOCUS       AR036196                111 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5871996.
ACCESSION   AR036196
VERSION     AR036196.1  GI:5952864
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Retrovirus agents MSRV1 and MSRV2 associated with multiple
            sclerosis
  JOURNAL   Patent: US 5871996-A 7 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1189
AR036716
LOCUS       AR036716                111 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5800980.
ACCESSION   AR036716
VERSION     AR036716.1  GI:5954572
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Detection of MSRV1 virus and MSRV2 pathogen and/or infective agent
            associated with multiple sclerosis, by nucleic acid hybridization
  JOURNAL   Patent: US 5800980-A 7 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|:::::|  |:|||:::||||:||||  |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1190
AR076978
LOCUS       AR076978                111 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 7 from patent US 5962217.
ACCESSION   AR076978
VERSION     AR076978.1  GI:10003724
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     MSRV1 virus and MSRV2 pathogen and/or infective agent associated
            with multiple sclerosis, and biopolymer constituents thereof
  JOURNAL   Patent: US 5962217-A 7 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
```

```
            :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|. |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1191
AR094699
LOCUS       AR094699                 111 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 7 from patent US 6001987.
ACCESSION   AR094699
VERSION     AR094699.1  GI:10021863
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with Multiple sclerosis
  JOURNAL   Patent: US 6001987-A 7 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1192
AR183934
LOCUS       AR183934                 111 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 7 from patent US 6342383.
ACCESSION   AR183934
VERSION     AR183934.1  GI:20227903
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Mallet,F., Mandrand,B., Bedin,F. and Beseme,F.
  TITLE     Cell lines and viral isolates associated with multiple sclerosis
  JOURNAL   Patent: US 6342383-A 7 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="unassigned DNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1193
AR343146
LOCUS       AR343146                 111 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 7 from patent US 6579526.
ACCESSION   AR343146
VERSION     AR343146.1  GI:33738648
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Viral material and nucleotide fragments associated with multiple
            sclerosis, for diagnostic, prophylactic and therapeutic purposes
  JOURNAL   Patent: US 6579526-A 7 17-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


RESULT 1194
AR344303
LOCUS       AR344303                 111 bp    mRNA    linear   PAT 17-AUG-2003
DEFINITION  Sequence 7 from patent US 6582703.
ACCESSION   AR344303
VERSION     AR344303.1  GI:33740244
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
            Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.
  TITLE     Isolated nucleotide sequences associated with multiple sclerosis or
            rheumatoid arthritis and a process of detecting
  JOURNAL   Patent: US 6582703-A 7 24-JUN-2003;
            Bio Merieux; Marcy l'Etoile;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="unknown"
                     /mol_type="mRNA"

  Query Match             1.1%;  Score 79.2;  DB 1;  Length 111;
  Best Local Similarity   58.2%;  Pred. No. 0;
  Matches   64;  Conservative   36;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2892 RTTTTACCCCAAGGGTTCAGGGATAGYCCCCATCTATTTGGCCAGGCATTAGCCCAAGAC 2951
             :| || || || ||||| |::|||| ::|:|||||:|||||:|::|:|:|::|:|:|:|:
Db          1 GTGTTGCCACAGGGGTTTARRGATANCYCYCATCTMTTTGGYCWRGYAYTRRCYCRAKAY 60

Qy       2952 TTGAGYCARTYMTCATACCTGGACACTCTTGTCCTTCRGTAKGTGGATGA 3001
             :|::||||:|::|:|::::::|  |:|||:::||||:|||| |||||||
Db         61 YTRRGYCAVTTCTYAKRYSYRGSNAYTCTBKYCCTTYRGTACATGGATGA 110


Search completed: August 8, 2006, 09:16:14
```

Job time : 3204 secs